Periodic Report (PR2)

Marc-Elian Bégin, Christophe Blanchet, Kathryn Cassidy, Evangelos Floros, Javier Fontan, Eduardo Huedo, Stuart Kenny, Ignacio Llorente, Charles Loomis, Louise Merifield, et al.

Enhancing Grid Infrastructures with
Virtualization and Cloud Technologies

Project Periodic Report

Periodic Report PR2 (V1.0)
15 June 2012

Grant Agreement Number	INFSO-RI-261552
Project acronym	StratusLab
Project title	Enhancing Grid Infrastructures with Virtualization and Cloud Technologies
Funding Scheme	CP/CSA
Date of latest version of Annex I against which the assessment will be made	2010-05-31
Periodic report	2nd
Period covered	2011-06-01 to 2012-05-31
Name, title and organisation of the scientific representative of the project's coordinator	Dr. Charles Loomis, Research Engineer, Centre National de la Recherche Scientifique (CNRS)
Tel	+33 (0)1 64 46 89 10
Fax	+33 (0)1 69 07 94 04
E-mail	loomis@lal.in2p3.fr
Project website address	http://stratuslab.eu

StratusLab is co-funded by the European Community's Seventh Framework Programme (Capacities) Grant Agreement INFSO-RI-261552.

Contributors

Name	Partner	Sections
Marc-Elian Bégin	SixSq	WP3, WP4
Christophe Blanchet	CNRS-IBCP	WP2, WP3
Kathryn Cassidy	TCD	WP3
Vangelis Floros	GRNET	WP3, WP5, WP6
Javier Fontan	UCM	WP4
Eduardo Huedo	UCM	WP6
Stuart Kenny	TCD	WP5
Ignacio M. Llorente	UCM	WP3
Charles Loomis	CNRS-LAL	WP2, WP3, WP5, Summary, Mgt.
Louise Merifield	SixSq	WP4
Ruben S. Montero	UCM	WP6
Henar Muñoz	TID	WP2, WP4, WP6
David O'Callaghan	TCD	WP3

Document History

Version	Date	Comment
0.1	15 May 2012	Initial outline.
0.2	7 June 2012	Complete draft for review.
0.9	13 June 2012	Proposed final draft.
1.0	15 June 2012	Final draft.

Contents

List of Figures

List of Tables

1 Publishable Summary

1.1 Project Context and Objectives

The StratusLab project is aimed at service provisioning, networking, and research of technologies that will bridge cloud and grid infrastructures to simplify and optimize the use and operation of existing distributed computing infrastructures (e.g. European Grid Infrastructure) and to provide a more flexible, dynamic computing environment for scientists.

The European production grid infrastructure has had many notable successes. It has allowed scientists from all over Europe and indeed from all over the world to federate their computing resources to advance their scientific aims. More importantly, the infrastructure allows them to federate their data and expertise to accomplish more than they would be able to do singlehandedly. Common APIs and service interfaces make it possible to take advantage of these distributed resources without having to modify applications for each site.

Despite its success, the grid also has its limitations. The uniformity of service interfaces unfortunately does not extend to the underlying computing resources, where users are exposed to significant heterogeneities in the computing environment, complicating applications and increasing failure rates. Passive calculations are handled well by the grid, but many applications require active services to coordinate the distributed analyses. Either scientists must provide their own resources for such services or negotiate with a particular site to provide them. This reduces the speed at which new calculations can be done.

Virtualization technologies provide a mechanism for offering customized, uniform environments for users with negligible performance degradation. Using grid technologies combined with virtualization allows the grid to provide users with a homogeneous computing environment, simplifying applications and reducing failures. Emerging cloud technologies allow users to dynamically allocate computing resources (often in less than a minute) and to specify the characteristics for the allocated resources. The fusion of cloud and grid technologies provides a more dynamic and flexible computing environment for grid application developers.

Cloud and virtualization technologies also offer other benefits to administrators of resource centers, such as the migration of live services for load balancing or the deployment of redundant servers. Reduced costs for managing resources immediately benefit users by freeing money for additional computing resources or

by having better user support from administrators.

A combined computing infrastructure that uses grid technology's strengths for federating resources, virtualization's strengths in providing custom, uniform environments, and the cloud's strengths in dynamic resource allocation, maximizes the utility of European distributed computing resources to scientists.

The StratusLab project creates an complete, coherent, open-source cloud distribution (for public, community or private clouds) to allow administrators of grid resources centers to take advantage of virtualization and cloud technologies. It provides new ways of using existing distributed computing resources to make the infrastructure more adaptable and more useful for scientists.

1.2 Summary of Work Performed and Achievements

The primary goal of the project was to create a complete, open-source IaaS cloud distribution, capable of running complex, production services like those found on the grid. This was achieved already in the first year of the project; the second year used to improve and to expand the cloud services while demonstrating the broad utility of cloud technologies with a variety of different academic and commercial applications.

The software is one concrete result of the StratusLab project, but other results have come from the operation of cloud services and from having worked together in a distributed collaboration. The following paragraphs highlight the most important results.

StratusLab Cloud Distribution The project has continued develop and enhance its IaaS (Infrastructure as a Service) cloud distribution. The distribution contains all of the expected services of a IaaS cloud: network, storage, and virtual machine management. Moreover, it provides innovative features like the Marketplace that facilitates sharing of virtual appliances, service management that allows deployment and autoscaling of multi-machine services, and support for multi-cloud scenarios. The distribution supports multiple operating systems (CentOS 6.2, Fedora 16, and OpenSuSE 12.1) and is ideal for both public and private cloud deployments.

Agile Practices The ability of the project to create two major releases of the project's software and approximately eight beta releases, in just two years, results largely from the project's adoption of agile software development methodologies, principally Scrum. The validation of these agile practices in a distributed, multinational collaboration is an important outcome of the project and provides ample justification for use of these practices in other, similar projects.

Reference Infrastructures Although principally oriented towards software development, the project created and maintained two reference cloud infrastructures, allowing scientists and engineers interested in cloud technologies to investigate them without having to go to the effort of installing their own cloud. Use of these infrastructures provided valuable feedback to the developers. More importantly

however, they allowed the administrators to develop knowhow and best practices for cloud operation. Both reference cloud infrastructures will continue after the end of the project, providing the kernels for larger national and international cloud infrastructures.

Adoption of the StratusLab Distribution A major success has been demonstrating the wide applicability of the StratusLab cloud technologies to scientific users, commercial users, and resource providers. The diverse scientific applications taking advantage of StratusLab include astrophysics, machine learning, software engineering, high energy physics, meterology, and bioinformatics. There are a number of StratusLab deployments by partner and non-partner institutes, including several commercial deployments. EGI is also evaluating the StratusLab Marketplace as a mechanism for managing and sharing virtual appliances.

Grid Services The project maintained a production grid resource center running over the StratusLab reference cloud infrastructure for nearly the entire duration of the project. This exercise demonstrated the maturity of cloud technologies and of the StratusLab cloud distribution. The ability to host complex, multi-machine services on a IaaS cloud with excellent performance validates the project's layered site deployment model, with the cloud providing an intermediate layer between the physical hardware and production services.

Multi-Cloud An important feature of the latest release is the support for multi-Cloud scenarios. Academic and commercial entities will naturally convert their existing computing infrastructures into clouds as they adopt cloud technologies. This means that in the short to medium term, users will be confronted with a "multi-cloud" computing landscape. The latest release allows users to take advantage of multiple clouds through bursting to public cloud providers and federation of peer cloud infrastructurs.

Cloud Consortium The project partners intend to form an open consortium to continue to evolve the StratusLab cloud distribution and to provide support for its users. Although the exact form of the consortium still needs to be defined, the first steps in evolving StratusLab from a project model to a community development model are already taking place. For example all of the documents and code have been moved to permanent repository to ensure access and to allow outside contributions.

1.3 Final Results and Potential Impact and Use

Most scientific and engineering research requires significant computing resources. Distributed computing infrastructures have brought unprecedented computational power to a wide range of scientific domains. Although, these architectures and the related software tools have been considerably improved over the years, they exhibit several difficulties, mainly due to limitations of physical platforms, which discourage adoption of grid technologies. StratusLab has the potential to profoundly change existing e-science infrastructures.

1.3.1 Improved Interdisciplinary Scientific Collaboration

Cloud technologies are expected to have significant impact, both immediate and long-term, in the way scientific research is carried out. Grid infrastructures have provided a remarkable advantage over the past years offering access to vast amount of computing power and storage space, and most importantly by offering a sustainable platform for scientific collaboration enabling the sharing of computing resources and scientific data. Cloud computing is expected to take this one step further by facilitating the easy deployment of customized grid infrastructures. These infrastructures are expected to have further positive impact on the way interdisciplinary scientific research is taking place.

StratusLab focuses on the provision of scientific infrastructures over cloud computing, investigating in particular the provision of customized Virtual Machine images. This customization will be done on the user side, which means that the user can have more immediate influence on the infrastructure itself. In this way the infrastructure will adapt to the user requirements and not vice-versa. By easing the management of grid sites and the configuration of hosting services we expect to attract a broader number of scientific communities and further facilitate their collaboration.

1.3.2 Impact on DCI Evolution

Currently, there is a big shift in all e-Infrastructure projects, and related efforts in Europe, to expand their activities in order to include cloud computing technologies. StratusLab will play a key role in this landscape by providing a focused environment for development, deployment and experimentation of cloud computing services.

The project provides an evolutionary path from the existing large-scale monolithic grid e-Infrastructures to novel, beyond the state-of-the-art, cloud-based, grid-enabled ones. Through its expected collaborations with other projects, StratusLab will disseminate its findings and drive direct impact on the way e-Infrastructure provision is currently done.

1.3.3 Improved Usability of DCI Platforms

Virtualization is the cornerstone of cloud computing and a key for achieving optimal usability of DCI platforms. Moreover, virtualized environments have the ability to adapt to different hardware platforms enabling a quick transition from one environment to another.

StratusLab operates such a virtualized platform on a variety of hardware environments. By offering customized machine images, users will be able to set-up an environment that better suits their application requirements. This will dramatically improve the current situation where current infrastructures are forced to offer a common configuration–a common denominator–that tries to do its best to satisfy many users with different runtime requirements. Another aspect where StratusLab will contribute is on power consumption efficiency (Green Computing) and the

Table 1.1: *StratusLab Information and Support*

Website	http://stratuslab.eu/
Twitter	@StratusLab
YouTube	http://www.youtube.com/user/StratusLab
Support	support@stratuslab.eu

Table 1.2: *StratusLab Partners*

CNRS	Centre National de la Recherche Scientifique	Charles LOOMIS loomis@lal.in2p3.fr
UCM	Universidad Complutense de Madrid	Ignacio LLORENTE llorente@dacya.ucm.es
GRNET	Greek Research and Technology Network S.A.	Evangelos FLOROS efloros@grnet.gr
SIXSQ	SixSq Sàrl	Marc-Elian BEGIN meb@sixsq.com
TID	Telefónica Investigación y Desarrollo SA	Henar MUÑOZ henar@tid.es
TCD	The Provost Fellows and Scholars of the College of the Holy and Undivided Trinity of Queen Elizabeth Near Dublin	David O'Callaghan david.ocallaghan@cs.tcd.ie

increased reliability by incorporating failover mechanisms using virtual machine snapshots and migration.

1.4 Contact Information

More information about the StratusLab project can be obtained from the sources listed in Table 1.1. Individual partners can also be contacted to obtain more specific information about their contributions to the project. Table 1.2 contains the list of StratusLab partners and relevant contacts.

2 Project Objectives for the Period

2.1 Objectives

The primary objective of the project is to provide a software distribution that brings together cloud and grid technologies to benefit both grid resource center administrators and scientists. In order to achieve this main objective, we have defined a set of interrelated objectives to be addressed in the project. The objectives are organized, for clarity of exposition, into three groups of objectives, corresponding to networking, service and research activities (see Figure 2.1):

- The first group represents coordination and networking with users and other stakeholders in the grid and cloud ecosystems. The project will work directly with scientists using the grid to ensure that the distribution satisfies real needs; and will collaborate with related projects and contribute to standards bodies.
- The second group represents infrastructure related services to the scientific community. The project will integrate and maintain a software distribution to bring cloud to existing and new grid sites and will ensure the production quality of the distribution by running two production sites with the distribution.
- The last group represents innovation and exploration of new cloud and virtualization technologies to enhance grid infrastructures. The project will develop innovative technology for cloud-like management of grid services and resources that will be incorporated into the software distribution.

These objectives are presented by work package below. Similarly, the work program is built around these objectives. There is a one-one correspondence between objectives and activities, so facilitating an easy cross-reference between objectives and activities throughout this document, and their verification during the project execution. The activity on project coordination has not been included here.

2.1.1 WP2: Interaction with Users and Related Communities

StratusLab targets two distinct communities: resource providers and end-users. The StratusLab software will simplify grid site administration and improve the reliability of the site. Later releases in the second phase of the project will provide

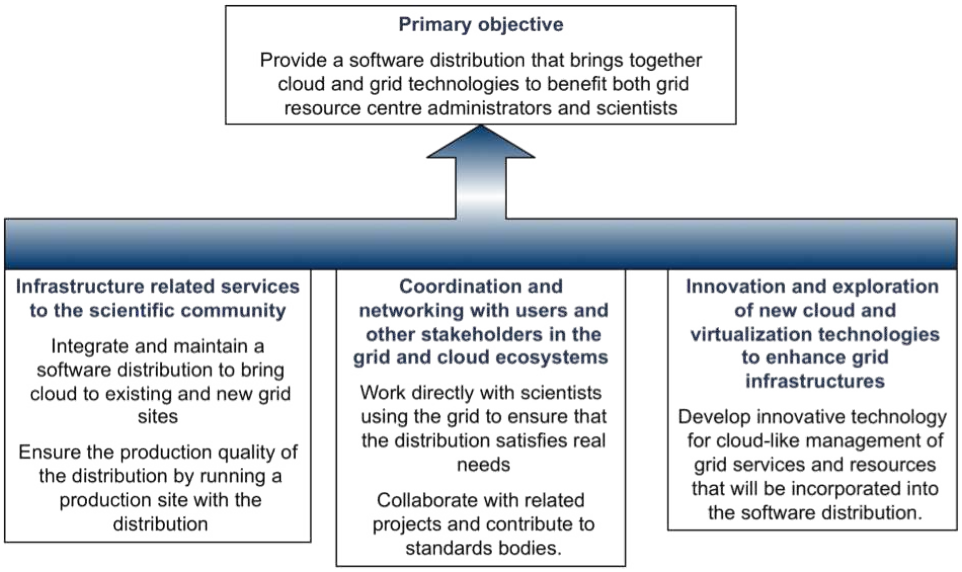

Figure 2.1: *Primary and supporting objectives*

direct cloud APIs that will be attractive for scientists porting applications to the grid. The communications between these communities and the project must be managed to ensure the project fully addresses their needs and any problems that arise. One community will work directly with the project to evaluate early releases of the software. Results of the project must be disseminated as widely as possible to those two communities as well as the general public. Scope of the objective.

- Manage communication with resource providers regarding their needs concerning virtualization and cloud technologies and their feedback on StratusLab software.
- Manage communication with end-users regarding their use of resources running StratusLab software and their needs for direct access to virtualization and cloud features.
- Training sessions will be organized to encourage dissemination of technical information and adoption of the StratusLab software.
- Evaluate early versions of StratusLab software from a users perspective with respect to utility and stability.

2.1.2 WP3: Dissemination

A large number of projects, companies, and standards bodies currently focus on cloud and virtualization technologies because of their promise and growing adoption. StratusLab must actively engage with those entities to ensure that the projects

results are well represented, that we are aware of others advances, and that we drive standardization in a direction consistent with our vision. Scope of the objective.

- Disseminate results of the project to resource providers, end-users, and the general public.
- Identify project contributions to standards bodies and standardization efforts.
- Coordinate interactions with related projects, developing Memoranda of Understanding between projects where appropriate.

2.1.3 WP4: Integration, Distribution and Support of Open-Source Cloud Distribution

StratusLab will integrate and support an open-source cloud distribution enabling grid site virtualization and dynamic scaling to remote clouds. This distribution will address the specific requirements of the grid resource providers and enable the deployment of science clouds, as well as addressing infrastructure cloud-like access demands from user communities, including industrial users. Scope of the objective. StratusLab will address the following topics:

- Selection of software components, from best of breed in open source software, to compose a robust and industry grade open source StratusLab toolkit. This distribution will integrate with typical administration tools in grid infrastructures and fabric management. This process will be driven by real needs and constraints in production infrastructures.
- Integration and management of open-source distribution, definition and maintenance of reference configurations and sustainability in the context of EGI and its official middleware distribution. The StratusLab toolkit will integrate the innovation developed in the research activity.
- Technical support for installation and configuration of the distribution, following industrial practices in term of quality, maintainability, testability and usability
- Definition of a process for automatic configuration of the virtual appliances

2.1.4 WP5: Operation of a Production Grid Site Running StratusLab

StratusLab will engage two resource centers that will be responsible for the deployment of middleware and tools developed in the project. One the main tasks of these resource centers will be the operation of two production grid sites running StratusLab toolkit. The sites should be able to pass the certification procedures imposed by EGI. The activity will demonstrate the security, performance, reliability and scalability of the distribution, and will provide support for the creation of

the virtual appliances for different user communities. The activity will also investigate the feasibility of offering a repository of reference images for cloud users, with demonstrated interoperability among the supported cloud infrastructures (including the private cloud deployed in the re-source centers, as well as a selected number of public clouds). Scope of the objective. StratusLab will address the following topics:

- Deployment and operation of virtualized grid sites
- Testbed for the StratusLab toolkit
- Support for the creation of virtual appliances for different user communities.

2.1.5 WP6: Innovative Cloud-like Management of Grid Services and Resources

StratusLab will conduct research on grid service automatic deployment and dynamic provision, including automatic elasticity mechanisms for scaling up and down to meet performance goals (typically defined by SLAs). StratusLab will also conduct research on novel infrastructure cloud-like resource provisioning paradigms, and dynamic and scalable management of virtualized infrastructures for grid services. The research will be performed to address technology gaps defined by the service activities according to user requirements collected by the networking activities. Scope of the objective. StratusLab will address the following topics:

- Framework for grid service elasticity and dynamic pro-vision of grid services
- Grid specific virtual machine management techniques
- Infrastructure cloud interfaces for grid sites and its integration with existing Grid services

2.2 Detailed Objectives by Quarter

2.2.1 Quarter 5

- Solidify the v1.0 StratusLab cloud distribution through increased testing and hardening of existing services.
- Support for a second operating system to ensure the portability of the distribution.
- Survey of the users and system administrators to see if the requirements have evolved from those already collected in Y1.
- Update and expand the target reference architecture for the distribution.
- Continued dissemination of project results.

- Continued operation of reference infrastructure and support to users and system administrators.
- Expansion of the number of users and sites using StratusLab.

2.2.2 Quarter 6

- Release incremental production versions of the StratusLab cloud distribution.
- Support for a second operating system to ensure the portability of the distribution.
- Implementation of an identified use case.
- Provision of tutorials for finding and training new users.
- Update and expand the target reference architecture for the distribution.
- Dissemination of project results with emphasis on the general public.
- Definition of reference cloud configurations and implementation of performance benchmarks.
- Improved integration of image management and caching.
- Implementation of more advanced networking services (e.g. dynamic firewalls).

2.2.3 Quarter 7

- Solidify the v1.0 StratusLab cloud distribution through increased testing and hardening of existing services.
- Support for a second operating system to ensure the portability of the distribution.
- Produce initial beta release of v2.0 of the StratusLab cloud distribution.
- Dissemination of project results of a completed use case.
- Expanding the sites using StratusLab and completion of MS4.
- Continued operation of reference infrastructure and support to users and system administrators.
- Operate a pre-production elastic Grid site elasticity, verify the applicability of the technology and move results to a production site.
- Evaluate GPFS as a backend storage solution. Prioritize and evaluate additional file systems.

- Develop additional use cases similar to MapReduce (e.g. Matlab application showcase)
- Integrate caching sub-system within the production cloud service
- Integrate NFS persistent storage service in the reference cloud service
- Integration of OpenNebula 3.0 with initial tests of new networking functionality

2.2.4 Quarter 8

- Release of v2.0 of the StratusLab distribution for multiple operating systems
- Detailed definition and initial implementation of sustainability plan
- Demonstrate of hybrid cloud computing
- Measurement of the behavior of the system in terms of scalability and IO performance, including with alternate configurations (e.g. GPFS)
- Demonstrate a pre-production elastic Grid site
- Complete migration of gLite production Grid site to UMD-1
- Finalise sustainability plans to ensure continued development and maintenance of project outputs

2.3 Review Recommendations

The responses given below to the reviewer recommendations have been updated to reflect the situation at the end of the project.

1. *Due to unscheduled availability of a physical production infrastructure experiments, testing and debugging have been affected. This was caused by problematic financial issues at the start of the project. The project should make provisions that such events are better mitigated in the future.*

   (See next response.)

2. *Several WP have shown a "delayed" start due to various hiring issues at the beginning of the project. A better process should be put in place.*

   These delays were largely related to the difficulty of hiring new personnel in the summer months and to lengthy administrative procedures. Most of the partners compensated to some extent with effort from permanent staff. Currently, all of the partners are fully staffed and no further perturbations are expected. In general, the overall deviations from the planned effort was around 10%, although because of larger than expected travel expenses, more expensive personnel, etc., the budgets are roughly in line with the planned budgets.

3. *The dissemination work is focussed towards the more technically oriented (system administrators) communities rather than the scientific user groups who could benefit from StratusLab. This is acceptable for year 1 but the focus should shift to the scientific users in year 2.*

   The project agrees that this is a reasonable shift in focus for the dissemination work in year 2. The project proposed a demo for the EU Innovation Convention, which unfortunately was not accepted. A general video highlighting the benefits of cloud technologies and showing concrete examples of use was prepared. Although still technical, it provides a better introduction to the cloud and its benefits then previous material.

4. *The dissemination targets as mentioned by the related KPI metrics were not very ambitious. The project should establish more ambitious KPI metrics with respect to dissemination targets.*

   The project management along with the activity managers made slight modifications to the metrics (mainly for the Marketplace) and updated some of the second year targets.

   In particular for dissemination, the target for people on the project announcement list was raised to 75 (with 73 currently enrolled). The goal for the discussion site was set at 100 where only 16 are enrolled, highlighting a problem we have had in jump starting a community around the StratusLab distribution.

5. *A clear and simple demonstration showing the benefits rather than the technology itself should be produced. This benefits should be illustrated focusing on one or two clear use cases. The solutions should clearly demonstrate what the real benefits are (for Scientific Users / System administrators) of the StratusLab toolkit.*

   A major part of the D2.3 deliverable was to identify possible use cases and would serve as a good basis for demonstrating the benefits of cloud technologies and as good topics for focused dissemination efforts. As mentioned above, a general video was created that describes the project, expected benefits from cloud technologies, and gives concrete examples of real use of the StratusLab cloud. This is a start for better demonstrating the cloud benefits. Deliverable D2.4 provides further examples, which will be used as a basis for continued dissemination of the Stratuslab consortium's achievements after the project.

6. *Provide a clear map of the components of the toolkit. Which components are re-used, which are newly developed and which are adapted from existing components?*

   Deliverable D4.4 that provides the reference architecture for StratusLab 2.0 was structured to provide a separate description for each service. That description includes details on where the component was developed, external

dependencies, and interactions with other components. This should provide a clearer overview of the development activities of the project. In addition, more effort has been made in the deliverables to describe clearly the project's work with respect to individual components, especially in the WP4 and WP6 deliverables.

7. *The periodic report is in draft status. Please submit a final version.*

   The final version of the periodic report for year 1 has been submitted. The delay was due the difficulty in getting the necessary financial information from the partner's institutes during the summer months. Through the PMB, the partners have been advised that this information must be provided in a more timely manner for the following periods.

   Although obtaining quarterly updates of the financial information turned out not to be possible, all of the partners have made better efforts in the preparation of the second year periodic report and and no delays are expected in the delivery of the final periodic report.

8. *Knowing year 1 budget under spending, a new forecast of the planned spending for year 2 including a recovery plan for the current under spending should be presented to the EC as soon as possible. A reasonable deadline is one month after the receipt of this review report. It should take into account the reasons of year 1 under spending and should introduce measures that allow the effective implementation of the recommendations for year 2.*

   Getting complete budget information from the partner's administrations quarterly has proven to be extremely difficult. Nonetheless, all of the partners were aware of their underspending and made a concerted effort to correct the problem in the second year. Preliminary figures indicate that all partners have invested more effort in the second year and have significantly closed the gap between the actual and planned budgets. See the financial section in the periodic report for details.

9. *The Data Management layer should be improved. In particular, StratusLab should be able to use existing and robust parallel file-systems which have better scalability than NFS such as Panasas or GPFS.*

   This recommendation refers to the use of shared file systems to make machine images available to the various computational hosts of the cloud infrastructure. Alternate technologies such as iSCSI that do not rely on a shared file system have been investigated and used successfully. This formed the basis for the persistent disk service for the 1.x series of releases.

   Nonetheless, use of shared file systems provide a convenient alternative and provides an easy entry configuration. A shared file system option was put back into the persistent disk service for the 2.0 release. The performance of various shared file systems was compared, including GPFS, GlusterFS, and Ceph. See the D5.5 deliverable for details.

10. *Testing and benchmarking in WP5 should be more detailed including performance aspects.*

    In the second year of the project, a more systematic approach to testing and validation of the StratusLab software was put into place. This included a certification testbed to validate all candidate releases; tests on this infrastructure were largely automated through Hudson. Following a successful validation, a separate pre-production infrastructure was upgraded to further validate the release. Together, complete installations of new systems and upgrades of existing systems were tested. These tests extended the daily automated installation and test of the development release of the software.

    Relevant application-oriented benchmarks have existed since near the beginning of the project. Unfortunately, these were never used to collect systematic metrics on the performance of the system. Instead effort focused on performance problems noted by users and system administrators. These included the caching performance of the persistent disk services and I/O bandwidth via the network and to local disk. Both of these have been improved in the v2.0 release.

11. *More emphasis should be put on the Cloud API rather than the GRID.*

    In the second year, there was a stronger emphasis on pure cloud services rather than on grid services over cloud. The latest release of the StratusLab cloud distribution has support for OCCI, TCloud and Deltacloud.

12. *Although security issues are taken very seriously, privacy issues should be taken seriously as well. For instance, in case when a "closed" Grid infrastructure is complemented/bridged with an external public Cloud infrastructure when it is processing for instance medical sensitive information that can be relayed back to an individual person.*

    Our feeling is that StratusLab should remain a "neutral" carrier that allows users and administrators to implement their own mechanisms for ensuring privacy and confidentiality. Nonetheless, we see areas in which the services can improve, peoviding, for instance, better logging and auditing information that can complement user-level strategies for privacy and confidentiality. Some minor improvements were made in this area.

13. *The security incident as reported in Q3 should be analysed thoroughly and measurements should be taken to prevent this to happen again on the live production system.*

    These security incidents were taken seriously and analysed thoroughly. As a result of these incidents two additional features were added to the StratusLab distribution: 1) enhanced logging and 2) image policy enforcement. The enhanced logging makes it easier to trace the characteristics, ownership, and history of a particular machine image aiding forensic analysis. The

distribution now also contains a policy enforcement engine that allows system administrators to define what images are authorized to run on a given infrastructure. This policy enforcement mechanism is closely tied to the information provided in the Marketplace. Cloud administrators will still need to monitor the cloud for suspicious activity and take corrective actions as necessary. From the operations point of view (WP5) we will remain alerted for potential future security incidents and will be ready to respond quickly in coordination with other infrastructure operations teams (NGI NOC, EGI CSIRT team etc.).

14. *The project should clearly define a small number of use cases and focus the project towards delivering real value to these user communities, targeting system administrators as well as scientific users. As also indicated later in section 5 "Use and dissemination of foreground" the project should adopt an attitude of someone who tries to offer services to the market and must convince someone to spend funds for the services. This way the project can maximize impact and the work can contribute to sustainability.*

    The deliverable D2.3 defines a set of seven initial use cases. These were partially implemented over the course of the second year and used to evaluate the functionality of the StratusLab cloud distribution. The deliverable D2.4 provides information on the scientific and commercial adoption of StratusLab. Many of those were used as concrete examples in the general video that describes the benefits of cloud technology. These examples form a good basis for "marketing" of the StratusLab distribution and more needs to be done to capitalize on these successes.

15. *A person should be appointed in the Project Management Board that can help the project to move from technology towards real end user solutions and benefits. This person should also help the project to establish solid relationships with stakeholders outside the traditional high energy physics (HEP) community.*

    The project agrees that having a "user champion" within the project is a good idea. However, the PMB is not the proper body as it meets infrequently and isn't in direct contact with the daily technical advances of the project. In the end, a viable solution for incorporating an outside representative has not been found. Instead we have relied on input from the scientists, engineers, and administrators with whom the project already has contact. In addition, commercial contacts through the partners have proven invaluable for validating the project's software and indicating new directions for improvement.

16. *While it is clear what StratusLab could offer to the scientific community, the impact of StratusLab would be much bigger when the toolkit could also be used for users in the commercial world. Through collaboration with the Spanish TID private Cloud project, a large number of potential additional*

use cases (Telefonica's customers) could be developed and should be taken seriously, including dissemination towards other DCI projects.

Better marketing of the StratusLab distribution will include both commercial and scientific communities as described above. TID will be more involved in presenting the project's work to enterprises with several possible venues being considered.

Recently several commercial ventures have shown interest in the StratusLab software and have created demonstrators with it. This is a significant validation of the StratusLab concept for use within the commercial sector.

SixSq has done an extraordinary job in promoting StratusLab in the commercial sector. This work includes using StratusLab as a platform for SixSq's own SlipStream product, a commercial turnkey cloud solution targeted at SMEs in collaboration with IBM and Darest, a software engineering platform for automated deployment of SCOS-2000 with ESA/ESOC, and a large StratusLab deployment by Atos within the Helix Nebula initiative.

Based on the outcomes of these initiatives, StratusLab has a clear niche as an efficient, easy-to-install private cloud solution for small to medium-sized infrastructures. However, the work with Atos also shows StratusLab's potential on large infrastructures as well.

17. *Dissemination of the StratusLab Toolkit should become more marketing driven and should target both ICT press to reach potential industry users as well as to potential VOs beyond the current bioinformatics users. A demo centred on a use case could help.*

    As stated in other responses, we agree that the dissemination should become more marketing driven. Deliverable D2.3 has identified specific use cases and deliverable D2.4 contains concrete success stories from a wide range of scientific and engineering domains. Specific dissemination activities and possible demonstrations will be considered as each use case is sucessfully implemented.

18. *Future reports should contain less "fat" and should be more cripy and to the point. Executive summaries should be self-contained and should answer: (a) why should I read the deliverable, (b) the benefits for my company/organisation, (c) aspects addressed in this deliverable, (d) summary of recommendations/findings. The report should clearly describe if components are newly developed, improved or reused and integrated by StratusLab.*

    We have made an effort with the second year deliverables to be more concise and provide better summaries.

19. *Ensure the project periodic and final reports are available at least two weeks prior to the review meeting.*

The review for year 2 will be scheduled to ensure that the final periodic report is available at least two weeks prior to the review. The partners have all been very responsive concerning financial information for the second year. No delays are expected for the second periodic report or the for second year deliverables.

20. *On top of the internal reports, consider publishing for the general press, like a newspaper and possibly to organise a public demonstration for the less technical audience, e.g. at the European Parliament.*

    In August, a general paper was prepared that describes the goals of the project and the StratusLab distribution. Although still technical, this gives a good overview of the project. This and other project documents will be further generalized to appeal to a wider audience. A first step in this direction was a proposal for a booth at the EU Innovation Convention in Brussels, which was unfortunately not accepted. Nonetheless, work on general materials and demos for non-technical audiences continues.

3 Progress and Achievements

3.1 Summary of Progress by Quarter

3.1.1 Quarter 5

In the fifth quarter, the project worked to finish and to disseminate the first production version of the StratusLab cloud distribution (v1.0). This release contains all of the core functionality required for an IaaS cloud deployment. Following this successful release, the project prepared for its first periodic review which took place at the beginning of July. For the remainder of the summer, work continued to consolidate services within the StratusLab distribution and to plan the work for the coming year, in addition to the standard operation of the reference infrastructure and other services. Highlights of the activities are given below.

Definition of Use Cases The D2.3 deliverable "Survey of Targeted Communities Concerning StratusLab" contains a set of identified use cases that can demonstrate the benefits of cloud technologies for real users and can serve as subjects for dissemination activites. Seven use cases have been defined and will be scheduled for implementation over the next year.

General StratusLab Paper A general paper concerning StratusLab's goals and results has been prepared and will appear in the book "European Research Activities in Cloud Computing" in March 2012. This provides a good technical overview of the project and will be used for dissemination and training activities.

Dissemination for v1.0 The v1.0 release of the StratusLab cloud distribution was a critical milestone for the project, marking the first production release of our software. Articles were published by EGI.eu and in iSGTW and news of the release was carried by HPC in the Cloud and HPCwire.

Evolution of Distribution The distribution continues to evolve including more functionality and becoming more robust. Improvements to the distribution include better integration with the persistent storage, feedback from operations, and client tools for Claudia. Code for RESTful tests has been included in the Marketplace to allow for better testing of the service. Work on including OpenNebula v3.0 continues.

MapReduce To demonstrate the appliability of StratusLab to different use cases, Hadoop was deployed inside of a StratusLab cloud. An appliance was create to support this and is generally available through the Marketplace. SlipStream from SixSq was used to deploy the system, demonstrating the integration of commercial tools with the StratusLab platform as well as the features of StratusLab itself.

New Scaling Policies The ability for autoscaling deployments via Claudia has been expanded with a number of new scale down policies added. This allows better control of the resources used in service deployements.

Most of the detailed objectives for Q5 have been achieved. The update of requirements from users and system administrators was not very successful because of the low number of survey responses. Nonetheless, we believe that the extensive requirements collected at the start of the project are still valid and form a reasonable basis for guiding the StratusLab work. The other objective not achived in Q5 was the full support of a second operating system. This work will be pushed into Q6.

3.1.2 Quarter 6

In the sixth quarter, the project worked to enhance the existing production release and to update the architecture for the second major release due at the end of the project. Unfortunately, a major cooling failure at LAL seriously perturbed the build and test infrastructure and consequently also the ability of the project to make incremental releases of the StratusLab cloud distribution. Nonetheless, one incremental release was made that was used to support a number of tutorials. Progress was made on all of the quarter's objectives, although many were not fully completed because of the above problem. The points below summarize the progress made.

Incremental Releases of the StratusLab Cloud Distribution Version 1.1 of the StratusLab cloud distribution was made publicly available. This version was quickly installed on the project's reference cloud infrastructure and on the LAL's test infrastructure. Two additional planned incremental releases unfortunately did not take place because of the perturbations to the build and test infrastructure described above.

Implementation of Use Cases In the previous quarter, seven use cases have been defined on which to focus the project's porting and support efforts. Work on three of these–a bioinformatics application, a commercial 3-tier prototype service, and a software development PaaS–have significantly advanced, although not yet to the stage where they have concrete results and can be publicized. All of these are expected to provide dissemination opportunities in the next quarter.

Support for OpenSuSE Support for a second operating system by the StratusLab cloud distribution is desired for a couple of reasons. First, it demonstrates portability of the code and services. Second, support of OpenSuSE will allow us to test GPFS, which should have much better performance than NFS, as a shared file

system for distributing images on the cloud infrastructure. Although modifications to the build system have been made to facilitate this, support for OpenSuSE has not yet been demonstrated.

Tutorials and Training Materials A major effort and success for this quarter has been a series of well-attended tutorials. These took place as part of the EGI Technical Forum in Lyon, France; as part of the ACGRID 3 school in Hanoi, Vietnam; and as a standalone event in Orsay, France. Overall, around 90 people participated in these events. The materials developed for these tutorials will continue to be updated for future releases of the distribution.

Restructured Communication Channels The communication channels between the project and the targeted communities have been somewhat restructured. In October an additional "user announcements" (announcement only) mailing list was created, specifically for users of the StratusLab reference infrastructure to inform them of upcoming service outages, upgrades and other important news. A new public user forum was created and went live on 16 November, and has been running on a trial basis so far. The aim of the forum is to provide a contact point for StratusLab users and to build a community support mechanism. The group already has 12 members, and has seen seven messages posted so far in two topics. It will thus be officially launched in December 2011.

Updated Architecture The quarter began with a redefinition of the global architecture of the StratusLab cloud distribution. The major changes were the additions of monitoring services, accounting services, and an "inter-cloud connector". The designs of components for advanced services and functionalities have also been updated, providing a clear roadmap for the second year developments.

Improved Image Management A major integration effort centered on the persistent storage service. The image management code was significantly refactored to take advantage of this service and to allow efficient caching of virtual machine images. With these changes, the primary delay for starting images is now the latencies associated with the scheduling procedure. The "create image" feature has also been refactored to take advantage of this service, drastically reducing the time necessary to create a new image.

Streamlined Release Procedures In order to better streamline the release and certification of new software releases, new procedures have been put in place, supported by changed in our continuous integration system. This should allow future releases to be done "with the click of a button" with nearly all of the test and certification tasks done automatically.

Reference Configuration and Performance Benchmarks Although an important topic, little time was spent on this in Q6 because it was more important to get the build and test system working correctly and refactored to support more fluid release procedures.

Implementation of Advanced Networking Services Work was done to provide advanced networking services such as the dynamic provisioning of VLANs and configuration of firewalls. These features are part of the OpenNebula 3.0 release and is in the process of being integrated into the mainline StratusLab distribution. Integration problems on both the StratusLab and OpenNebula sides are being resolved and these new features will appear in the StratusLab distribution once OpenNebula 3.0 has been integrated.

Sustainability As we reach the end of the StratusLab project, sustainability is an important question for the users of our software. It is a question that comes up at nearly every StratusLab presentation. The face-to-face meeting at the end of the quarter allowed a detailed discussion of the sustainability strategy that will serve as a good basis for the final plans that will be developed over the last six months of the project.

Despite the problems with the build and test infrastructure's availability, significant progress was made on all of the objectives for this quarter. Those that have not been completed will be moved forward into the next quarter. More emphasis will be placed on supporting a second operating system as this has been pushed forward two quarters now. The next quarter will see the beta of the 2.0 release with prototypes of all of the expected services.

3.1.3 Quarter 7

In the seventh quarter, the project concentrated its efforts in four areas: improving the StratusLab distribution (through bug fixing and expanded testing), demonstrating cloud use cases with real applications, expanding the functionality of the services. The improvements and new functionality have been rolled into the incremental releases of the distribution. Two such releases (v1.2 and v1.3) have been made in this quarter, acting as beta releases for the upcoming v2.0 release at the end of the project.

Implementation of Use Cases Signficant progress has been made in running prototype and real applications on StratusLab cloud infrastructures. The most advanced are the two bioinformatics use cases–one providing bioinformatic web services and the other, TOSCANI, an application from Institut Pasteur. In addition a proof-of-concept multi-tier application has been prepared to demonstrate the applicability of the cloud to commercial use cases. Initial steps to interface StratusLab with the high-energy physics analysis framework DIRAC have been taken.

Improvements of Cloud Distribution External users from the above applications have identified many areas where the user interface can be improved. During this quarter many minor bugs have been corrected in the command line interface. As part of the bioinformatics efforts, a web-based graphical interface has been developed to provide customized access to the cloud services; this complements existing service web-based interfaces. As the pdisk server has been more closely integrated into the overall virtual machine workflow, problems with scalability and performance have been noted. Significant effort has gone into solving those issues.

Integration of OpenNebula 3.0 The integration of the more recent version of OpenNebula makes many of the new features available to users of the StratusLab distribution. These features include better network management and security features as well as more options for controlling placement of virtual machines. Integration of OpenNebula 3.2 is ongoing, but should happen quickly as there are fewer fundamental changes between these releases.

Improved Testing The releases have also been hardened with more thorough testing of existing features. The certification procedure has been improved to help ensure that fewer problems appear in production. Moreover, the procedure has been automated to make the procedure quicker and more efficient. This will further improve in the future with the possibility of scalability testing of StratusLab on the Grid'5000 infrastructure and the validation of the release on multiple operating systems.

Expanded Functionality Development of functionality intended for the final releases of the StratusLab distribution is already well advanced. Services to provide machine and application-level monitoring have been developed and integrated with Claudia, allowing more advanced autoscaling capabilities. In addition, some work has been done on standard service interfaces/libraries, including TCloud, OCCI, and DeltaCloud.

Sustainability Detailed planning for sustaining the StratusLab developments after the project started in the previous quarter and have been further refined through participation in a Sustainability Workshop between the DCI projects in January and collaboration with the ERINA+ project. This ERINA+ collaboration will, in particular, allow us to better understand the socioeconomic impact of the project and better define the stakeholders involved.

Despite these achievements, several problems have arisen in this quarter and consequently some of the quarter's objectives have not been met. The most notable problem was with the performance and scaling issues coming from the storage service; these problems appeared as the storage service was more tightly integrated into the virtual machine management workflow. Significant effort was put into understanding and resolving these issues as quickly as possible. Consequently, the upcoming 1.4 release will provide a much improved storage service, although there are still further improvements to be made. Because of this change of focus, tests of other storage services like GPFS have not been done.

On the application side, signficant progress was made in running real applications on the cloud infrastructure. Unfortunately, these successes have not been adequately publicized. Presentations at the EGI Community Forum early in Q8 will compensate somewhat for this, but a separate special effort needs to be made to maximize our benefits from these activities.

Overall, the StratusLab distribution is advancing well with the project correcting any problems that arise and adding new functionality at it moves toward our v2.0 release in the next and final quarter of the project.

3.1.4 Quarter 8

In the eighth and final quarter, the project concentrated its efforts on the finalization and release of v2.0, demonstrating the functionality via scientific and commercial applications, and working towards the implementation of the defined sustainability plan. The overall achievements of the project are contained in a large number of final deliverables covering all aspects of the project.

Adoption of the StratusLab Distribution A major activity of this final quarter has been demonstrating the wide applicability of the StratusLab cloud technologies to scientific users, commercial users, and resource providers. The diverse scientific applications taking advantage of StratusLab include astrophysics, machine learning, software engineering, high energy physics, meterology, and bioinformatics. There are a number of StratusLab deployments by partner and non-partner institutes, including several commercial deployments. EGI is also evaluating the StratusLab Marketplace as a mechanism for managing and sharing virtual appliances. The broad utility of the project's software in both academic and commercial settings bodes well for continued demand after the project.

Sustainability Plan A final plan for exploitation and sustainability has been put in place which allows for the continued development and maintenance of the StratusLab Distribution past the end of the project. The plan foresees a transition to an open, community-based consortium. Steps have already been taken in this direction to ensure continued availablility of the project's software and a smooth transition from project to consortium. For example, all of the software is now hosted in GitHub and all documents are available in an open-access repository.

StratusLab Cloud Distribution v2.0 The highlight of this last quarter is the finalization and release of the second major version of the project's cloud distribution. This release consolidates the functionality enhancements and bug fixes from the past year, integrates OpenNebula 3.2, demonstrates multi-cloud scenarios, enhanced Marketplace functionality, and support for IPv6. The release supports multiple operating systems: CentOS 6.2, Fedora 16, and OpenSuSE 12.1. As for previous releases, a concerted dissemination effort will make people in the wider scientific and technical communities aware of this release.

Reference Infrastructures The reference cloud infrastructures have followed the incremental releases of the StratusLab distribution, fine tuning the services to provide the best platforms possible for the StratusLab users and providing feedback (especially on storage services) to the developers. A second infrastructure was opened to the public at CNRS/LAL to complement the existing infrastructure at GRNET. A common authentication service was put in place between the infrastructures to allow users to have access to more resources and to allow them to test cloud federation strategies.

Both CNRS/LAL and GRNET will continue to operate their cloud infrastructures as a service after the end of the project. Providing demonstration platforms for the nascent open-source consortium and ensuring that current users do not see

a disruption of service.

Multi-Cloud An important feature of release v2.0 is the support for multi-Cloud scenarios. Claudia has been enhanced with a placement module allowing resources in different clouds to be used transparently in a cloud brokering scenario. Similarly, OpenNebula has been enhanced to allow cloud bursting to public clouds and federation with other StratusLab infrastructures.

Overall, the project participants have broadly achieved the goal of the StratusLab project: creating a complete, open-source distribution for an Infrastructure-as-a-Service cloud. The relevance of the provided services has been shown via the production grid site running over a StratusLab cloud and by the wide variety of scientific and commercial applications using StratusLab. This forms a solid foundation for the transition from a project to an open-source community consortium.

3.2 WP2: Interaction with Targeted Communities

This activity manages the relationships with the communities targeted by the project, notably scientists from diverse fields and system administrators interested in deploying a cloud infrastructure. Through those interactions, it provides requirements and feedback to the other activities within the project while at the same time evaluating the StratusLab distribution from the points-of-view of users and system administrators. The activity also helps provide support to the targeted communities.

3.2.1 Summary

3.2.1.1 Quarter 5

After preparations for the first EU review of the project, the primary activity in this quarter has been the work on the deliverable D2.3 "Survey of Targeted Communities Concerning StratusLab". This deliverable identified seven concrete use cases to be implemented during the second year of the project. Approximately half of them are "internal" use cases that can be implemented using only StratusLab personnel; the other half involve interaction with scientists and engineers outside of the project. Once implemented, the results will be disseminated to the wider community (in cooperation with WP3) to encourage broader usage of the StratusLab cloud distribution.

This deliverable represents the first action in refocusing the WP2 efforts to bring more users onto StratusLab cloud infrastructures. This was the principal recommendation for WP2 coming from the first European review in June. This refocusing of purpose will continue as concrete plans and schedules for the identified use cases are made.

3.2.1.2 Quarter 6

The major effort for this period has been giving tutorials for the StratusLab cloud distribution. Three large tutorials were given (Lyon, France; Hanoi, Vietnam, and Orsay, France) that included around 90 students. In addition, a cloud module featuring StratusLab has been added to an software engineering masters programs in Marrakech, Morrocco with 18 students. The developers of the Narwal data aquisition framework ran a tutorial of their software using the LAL StratusLab cloud infrastructure, allowing them to provide easily temporary resources for the students. This activity by an outside institute is an important validation of the StratusLab software. In addition to these training activities, the work package has worked on specific use cases (commercial and academic) to demonstrate the capabilities of the StratusLab cloud distribution.

3.2.1.3 Quarter 7

The work in this quarter has been concentrated in three areas: porting of applications, usability improvements, and expanding the tests and validation of the software. These applications include two from the bioinformatics domain (TOSCANI

and Bioinformatics Web Service) that allow people from outside the project evaluate the cloud paradigm and the StratusLab software for their scientific analyses. A prototype application using the standard 3-tier architecture and showing the benefits of autoscaling technique available in Claudia has been created and will be used as a concrete demonstrator of the project's software. Based on feedback from users, the command line interface has been improved; in addition, a web-based interface has been created for the bioinformatics users to simplify access. Also to improve the software, people in WP2 have concentrated on improving the functionality testing, laying the foundations for scalability testing (via the Grid'5000 platform), and providing packages for multiple operating systems to permit GPFS testing. These achievements will be presented at the EGI Community Forum at the end of March and then followed by more formal press releases.

3.2.1.4 Quarter 8

A major activity of this final quarter has been demonstrating the wide applicability of the StratusLab cloud technologies to scientific users, commercial users, and resource providers. The diverse scientific applications taking advantage of StratusLab include astrophysics, machine learning, software engineering, high energy physics, meterology, and bioinformatics. People from these areas were helped with direct support and training. There are a number of StratusLab deployments by partner and non-partner institutes, including several commercial deployments. Details are provided in the D2.4 deliverable. A companion deliverable (D2.5) provides an evaluation of the StratusLab distribution against defined requirements and feedback from users and administrators. The broad utility of the project's software bodes well for continued use as the partners transition to an open, community-based consortium.

3.2.2 Task 2.1: Interactions with Resource Providers and End-users

3.2.2.1 Quarter 5

Survey of Targeted Communities (D2.3) The primary activity of the last quarter has been the deliverable D2.3 "Survey of Targeted Communities Concerning StratusLab". In keeping with the recommendations coming from the first European review of the project, the focus of this deliverable shifted slightly to identify specific use cases that should be implemented during the second year of the project. Seven different use cases have been identified; three of which are "internal" use cases that can be done using only effort within the project and the remaining four involve scientists and engineers outside of the project. An attempt was also made to update the requirements collected in D2.1. However, this was largely unsuccessful because of the summer holidays. Instead the project will continue to rely on the requirements listed in D2.1 and feedback from people as we work to fulfill the identified use cases in D2.3.

General Overview of StratusLab A general overview of the project's goals, developments, and results has been prepared based on a presentation at the Second Workshop on Software Services (WoSS) in Romania. This overview will appear as a chapter in the book "European Research Activities in Cloud Computing" (Dana Petru and Jose Luis Vasquez Poletti, editors) that should appear in March 2012.

Bioinformatics Use Cases CNRS IBCP has participated to the "Survey of Targeted Communities Concerning StratusLab" described in the project Deliverable D2.3. We have proposed two use cases: one about Bioinformatics Web Services and the other, TOSCANI, dealing with the determination of protein structures based on Nuclear Magnetic Resonance (NMR) information. The goal of the first bioinformatics use case is to create bioinformatics appliances containing the usual tools that scientists and engineers can deploy on demand. To meet researcher's expectations, these appliances must present a standard programmatic, public, web service interface, permitting users to combine the different bioinformatics methods in useful analysis pipelines. The second bioinformatics use case has been defined in collaboration with the M. Nilges group at the Institut Pasteur Paris. TOSCANI, "TOwards StruCtural AssignmeNt Improvement" concerns the scientific disciplines around (i) molecular and structural biology (determination of biomolecular structures up to atomic resolution) and (ii) bioinformatics, which includes the ensemble of computer algorithms for treating the data from biological systems. We propose there to demonstrate the flexibility of the cloud to deploy the different bioinformatics tools required to improve the determination of protein structures based on NMR information.

Interaction with French Bioinformatics Community Members of IBCP have participated to a meeting of the RENABI GRISBI community on June 21-22 in Lyon. Different points were discussed about the distributed infrastructure GRISBI serving the national community and the reference bioinformatics data and tools that are deployed. This is designed to foster acceptance and deployment of the StratusLab Toolkit within this community and to gather their requirements.

Telefónica Private Cloud Contact Telefónica is offering a Private Cloud for the Spanish regional administration, where TID is developing part of this solution. In fact, they are developing the product from ideas and the service manager prototype, which is released in StratusLab. In the quarter, some contacts have been done with the product leader in order to obtain feedback with end-users of the service.

Commercial Use Case Identification In order to provide requirements and feedback from a commercial perspective, TID has been working on the identification of a commercial use case to be used in order to validate the StratusLab distribution and show their advantages.

3.2.2.2 Quarter 6

Tutorials People from this activity have worked intensively this quarter to update existing teaching materials and develop new material for recent releases of the Stra-

tusLab distribution. These materials were used in three tutorials, collectively with around 90 students. The first was at the EGI Technical forum in Lyon, France. The second was in Hanoi, Vietnam (1–3 November) as the cloud part of the ACGRID-III school. The third took place in Orsay, France (17–18 November). All tutorials used the LAL test cloud infrastucture. All were successful with good feedback from the students, although the Vietnam tutorial was affected significantly by a lack of resources because of air conditioning problems at LAL. The agendas for these tutorial are listed in the dissemination section.

Narval Tutorial From the 15-21 October 2011 there was the first "Narval School" organized by the Nuclear Physics Institute (IPN) in Orsay, France. Narval is a highly-distributed and modular framework for data acquisition. This tutorial took place on the StratusLab cloud infrastructure at LAL. The use of this infrastructure allowed the organizers to provide a flexible test infrastructure for the students with a minimum of effort. This tutorial, by an outside institute, was an important validation of the StratusLab software and shows the cloud's utility for supplying temporary infrastructures.

Cloud Module in Engineering Curriculum A new module on "Cloud Computing and StratusLab" was added to the Information System and Networking degree program at the Faculty of Science and Technology in Marrakech, Morrocco. This course has 18 students and a total duration of 24 hours. The goals of the course are to allow the students to identify the basic components, types, and modules of Cloud computing and understand how they function. StratusLab played a key role by providing a concrete set of IaaS services and an open infrastructure for experimentation.

CernVM CernVM is an initiative from CERN to provide virtual machine images adapted to the needs of the high-energy physics community. In particular, these images contain the software of the four Large Hadron Collider (LHC) experiments. LAL personnel have worked with the CernVM developers to allow these images to be used directly on a StratusLab cloud infrastructure. Differences in the contextualization procedure have been resolved (on both sides) to allow future versions of these images to run on StratusLab resources.

Alternate Schedulers The scheduling algorithm provided natively by OpenNebula is functional but quite basic. We have been approached by two groups with more elaborate schedulers that are interested to integrate them with the StratusLab distribution: first, to show that such integration is possible and second, to understand how they behave with real users. One group is from INRIA Lille Nord[1] in Lille, France; the other is based at Distributed Systems Research Lab (DSRL)[2] in Romania. Work continues with these groups on the integration.

E-business Applications Use Case An e-business application use case is being defined and implemented inside WP2 in order to validate the StratusLab distri-

[1] http://www.inria.fr/centre/lille/

[2] http://dsrl.coned.utcluj.ro/

bution and show its advantages. This e-business application use case involves the deployment of a typical enterprise application (composed by front-end, business logic and database) in the StratusLab infrastrastucture. Their main requirements are simplicity (automatic deployment), multi-tier management, scalability and security (by LAN and firewall). Four images have been developed and are registered in the StratusLab Marketplace.

Bioinformatics Web Service Use Case We have added new bioinformatics tools to the "biocompute" appliance and updated existing ones. We have also added mechanisms for mounting automatically biological databases repository at start-up. Users wanting to run a biocompute instance connected to the biodatabases repository of the cloud infrastructure simply set the appropriate options when launching the machine. For example on IBCP's cloud by setting the BIO_DB_SERVER parameter to the value 'idb-databases.ibcp.fr'. We have also instrumented the appliances to use the StatusLab persistent disk functionality.

TOSCANI Use Case TOSCANI deals with the determination of protein structures based on Nuclear Magnetic Resonance (NMR) information. In collaboration with the M. Nilges group at the Institut Pasteur Paris, we matched the ARIA requirements with the StratusLab services. The main requirements are the management of the input and ouput data (at the scale of GB), the deployment of a complete infrastructure, including one master and up to 100 compute nodes for usual experiments. Shared storage, between the ARIA master and CNSsolve compute nodes, is required; additionally, the data produced need to be kept for further analysis. The persistent disk utility corresponds well to this requirement.

Interaction with French Bioinformatics Community Members of IBCP have participated in a workshop of the RENABI GRISBI community on 2-4 November in Rennes. We have discussed the requirements about representative bioinformatics databases and tools in the context of the distributed infrastructure GRISBI serving the national community. We have also demonstrated the StratusLab facilities using IBCP's cloud infrastructure. This is intended to foster acceptance and deployment of the StratusLab cloud distribution within this community and to gather feedback concerning their requirements.

3.2.2.3 Quarter 7

E-business Application Use Case The e-business application use case defined in Q6 has been developed during this quarter. The four images required: a front-end, a back-end, a balancer and a data store, have been uploaded to the appliance repository[3] and made available through the Marketplace. In addition, a set of monitoring probes are been developed to monitor the number of users and response time in the application, Key Performance Indicators (KPIs) that will drive the scalability. Finally, an OVF description has been defined for the deployment of the use case, its monitoring and scalability.

[3] http://appliances.stratuslab.eu/images/comertial

Bioinformatics Web Service Use Case New releases of the 'biodata' and 'biocompute' appliances have been built and registered in the StratusLab Marketplace. The 'biodata' appliance now contains the SwissProt and PROSITE databases; bioinformaticians can use the appliance to set up their own cloud biological databases repository. They can add the databases they want with the help of the BioMaj system. The configuration of the databases is done through the BioMaj web interface or through the command line. This appliance can be used in conjunction with the 'biocompute' appliance to establish a biological repository for future bioinformatics analyses. The 'biocompute' appliance contains a range of tools (BLAST, ClustalW2, FastA, HMM, etc.). And the tool names can be used to apply a filter on the bioinformatics appliances at the "launch instance" step. The access to this appliance is also provided to users through SSH or through a web portal, easily accessible in the main window of the cloud web interface. Biocompute instances can mount the biological database repository of the local cloud (if available) through contextualization options. This appliance can also mount your persistent disk, at boot time if users choose via "launch instance" form.

TOSCANI Use Case We have released two versions of the ARIA appliance related to the TOSCANI use case. ARIA is developed by the group of Dr. M. Nilges at the Institut Pasteur Paris. Version 2.3 of the tool has been installed and configured, and the appliance has been registered in the Marketplace. The appliance has been evaluated by colleagues from Institut Pasteur group. They had run experiments with real data and reported that everything went fine, with good performances of the CNSsolve part. A second version of the appliance (1.1) has created with more computing (BLAS, Lapack) and OS (screen) add-ons. The user connects to the ARIA VM with SSH, uploads its data and can directly run the tool from the current directory. All the tools and libraries have been put in the environment variables. Only one instance, containing both the ARIA master and the CNSsolve tools, needs to be run by the user. The computing resources are tuned via the instance used to launch the machine, up to the maximum allowed by the "HTC" instance on the CNRS IBCP infrastructure with 24 CPUs and 16 GB RAM. Coming releases are planned which will allow an instance for the ARIA master and many instances for the CNSsolve nodes.

Specialized Cloud Web Interface for Bioinformatics CNRS IBCP has developed a new web interface for the StratusLab framework. The main goals are to make it easy to use by non-computing and non-cloud-specialist scientists like biologists and bioinformaticians, and powerful enough to perform the main tasks of the VM lifecycle on a StratusLab infrastructure. The first release has been developed and is now being evaluated by biologists and bioinformaticians. The available features allow the creation and termination of the virtual machines, allow the management of the persistent disks, and provide assistance for bioinformatics appliance selection and instance contextualization. A screenshot of the interface can be seen in Figure 3.1.

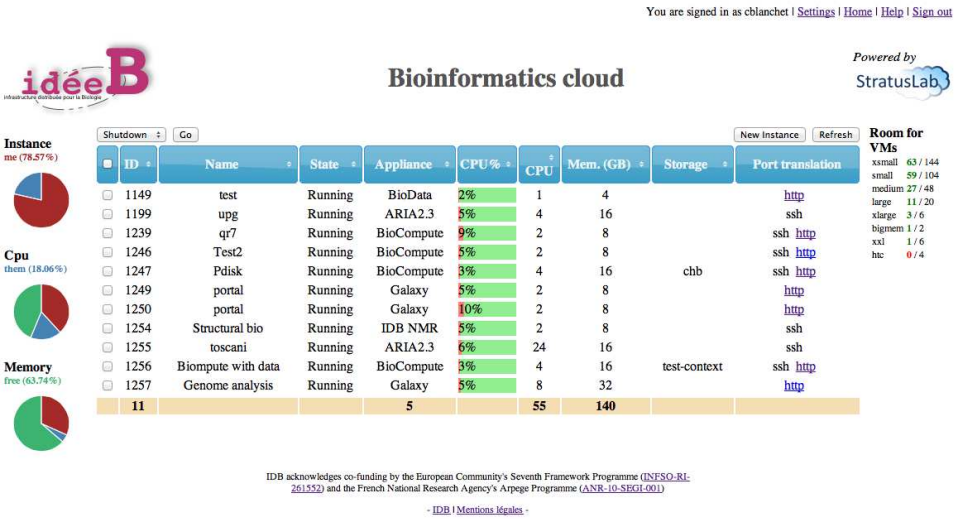

Figure 3.1: *Bioinformatics Cloud Web Interface*

DIRAC: HEP Use Case The high-energy physics community represents an important use case for the project because its current use of grid services offers the opportunity to understand the interactions between grid and cloud services. Modifications required to run CernVM appliances (machine images created by the community) have been completed and released. A test has been added to the standard "use case" Hudson job to ensure this continues to work with future versions. As a follow-on to this, work has begun with the developers of DIRAC, a master-worker analysis framework used by the LHCb experiment based at CERN. The goal is the transparent integration of StratusLab resources on the cloud within the DIRAC framework, with a demonstration of production jobs being run on our infrastructure. This integration should be trivial with the availability of the CernVM images and we expect concrete results early in Q8.

Automation of Base Image Production The project has provided minimal machine images for ttylinux, Fedora, and Ubuntu for some time. However, these images have not been systematically kept up-to-date or produced in a consistent manner. Now the mechanisms for producing these base images have been automated (through Hudson) that will allow systematic production and updates. In addition, the range of supported base images has been expanded to include also OpenSuSE and CentOS. This will allow users to find high-quality images that are guaranteed to work well on StratusLab cloud infrastructures. This work was done in collaboration with WP5.

Improvements to Client Based on feedback from users at the project's tutorials a number of usability issues were identified with the StratusLab command line client. During this period a number of these issues were resolved with the client. These improvements generally provide better error messages to users and improve

the overall consistency of the tools.

Multi-platform Support Support for multiple operating systems has been a long-standing request of system administrators. (The client already supports a large range of operating systems.) This has become more urgent as tests of the GPFS file system as an image sharing mechanism require support for a platform other than Fedora 14. Moreover, Fedora 14 no longer receives security updates, so system administrators are, quite reasonably, hesitant to install that system. During this quarter, the necessary jobs and resources were put in place to provide packages for all of the StratusLab services on Fedora 14, OpenSuSE 12.1, and CentOS 6.2. This will allow tests with GPFS to proceed and will give system administrators the ability to use operating systems with long-term support. The project will move to CentOS 6.2 as the core platform for the next release and will also upgrade to use Fedora 16.

3.2.2.4 Quarter 8

Tutorials WP2 to update the user tutorial to reflect the most recent releases. A tutorial was given at the EGI Community Forum in Munich, Germany with an attendance of around 25 people. The dissemination section lists this tutorial and provides a link to the agenda and material. This tutorial was used to give a "private" tutorial to four people from the Astroparticle and Cosmology (APC) laboratory in Paris, France. The results of their initial use of the StratusLab cloud are described in the deliverable D2.4.

Deliverable D2.4 The Deliverable D2.4 documents the adoption of the StratusLab distribution by end-users and system administrators. This provides a summary of our work with people in these two groups over the project's lifetime. It also provides retrospective of our activities with the users and system administrators and draws some general conclusions on the successful and less successful aspects of our work.

AppStat Researchers from the Machine Learning and Applied Statistics Group have started to use the StratusLab cloud infrastructure for their research into machine learning algorithms. The project support personnel from them port an initial application to the cloud and to use the resources for an initial scientific analysis. This activity is fully reported in the deliverable D2.4.

Updated Base Images The previous work done to automate the production of base machine images (i.e. appliances with minimal distributions of common operating systems) has allowed this images to be updated frequently with operating system patches and when problems were found with the images. These changes include increasing the disk space available in the images, ensuring that hostnames are not embedded in the images, and enabling IPv6.

Project Video A new video has been produced that gives an overview of the project, its software, and examples of use. This eight-minute video is available from YouTube and linked from the project website. This serves as a good intro-

duction to more detailed videos that will be produced for the review concerning bioinformatics and the n-tier application prototype.

Enhanced in the E-business Application Use Case The e-business application used for demonstrating advanced StratusLab functionalities developed in WP6 (multi-tier service management, scalability and federation) has been evolved providing its database stored in the persistent disk and a GUI.

Update of the Bioinformatics Service Use case In the context of Bioinformatics Web Service use case, CNRS IBCP have added new bioinformatics software to the 'biocompute' appliance. For example, new tools about multiple alignment of sequences (gene and proteins) like Muscle and Clustal Omega have been installed and configured. This appliances are still fully contextualized to be used with the 'biodata' appliance and the persistent disk feature.

Update of the TOSCANI Use Case A new version (2.0) of the ARIA2.3 appliance following the IaaS model has been developed and released in collaboration with the group of Dr. M. Nilges at the Institut Pasteur Paris. With the usual cloud interfaces (command-line client and Web interface on IBCP's infrastructure) scientists can launch multiple virtual machines that will be used for a compute-intensive analysis. They connect to the virtual machine that will be used as the ARIA master and run the script 'aria-clouder' we have developed and installed in the ARIA2.3 appliance. This tool finishes the configuration of the virtual cluster, setting up the system services and the ARIA configuration file to adapt to the virtual infrastructure. Scientists have then simply to run the ARIA tool as usual and retrieve their results once the computations finish. This appliances is also compatible with all the instance types available on IBCP's cloud (large, bigmem, HTC) and with the StratusLab reference infrastructures at GRNET and LAL.

Online Bioinformatics Documentation CNRS IBCP has updated the online bioinformatics documentation[4]. The scientific context is described to motivate the development we are doing in the project. The two bioinformatics use cases are described in detail there. Similarly, the three bioinformatics appliances are also described, as a complement to the Marketplace descriptions.

Usage of the IBCP's Cloud Infrastructure by the Bioinformatics Community The IBCP infrastructure has been opened other colleagues from the bioinformatics community. New accounts have been created and we have run experiments with real data for protein modeling. New contacts have been initiated with scientists working in the genomics area. This action will bring feedback from scientists to the project developers.

Bioinformatics Web Interface for StratusLab Cloud CNRS/IBCP has added new features to its bioinformatics web interface to the cloud. Scientists can now manage their persistent disks through the web interface. They can create as many disks as they need, within their defined quota. There also new features for cloud

[4] http://www.stratuslab.eu/doku.php/bioinfo

administrators to help manage the all users' disks easily.

The two bioinformatics use cases and the bioinformatics web interface for StratusLab cloud have been presented at the "Clouds: Users" session of the EGI Community Forum in Garching on March 2012.

3.2.3 Task 2.2: Intensive Evaluation of StratusLab Products

3.2.3.1 Quarter 5

Evaluation of StratusLab Installation in a representative Bioinformatics Lab
CNRS IBCP has evaluated the manual and the quattor-based installation modes of the StratusLab framework. Several issues have been identified and reported to the project's developers, who have corrected most of them. Deploying a cloud site on CNRS IBCP hardware resources aims at identifying site constraints usually encountered with national bioinformatics sites (not enough public IPs, NAT, etc.), to provide bioinformatics scientists and engineers with computing access and usage suitable to their customs (simple interface, Web if possible, Shibboleth credentials), and to deploy specific cloud services devoted to Bioinformatics. Some developments are being made to adapt the StratusLab components to these specific constraints: PAT forwarding (port address translation) to satisfy the lack of public IP addresses, "snapshot mode" to create a new virtual machine from a previous base appliance, evaluation of the integration of Shibboleth authentication procedure. Manual installation is potentially the most realistic way of deploying the StratusLab system in most bioinformatics laboratories that do not have enough system administrators. Quattor-based installation mode is also a good perspective for the sites of the French bioinformatics distributed infrastructure RENABI GRISBI who have begun to use the Quattor system these last years. The main goal of this evaluation was to fix bugs and to confirm the reliability of these procedures for the public release 1.0 of StratusLab.

3.2.3.2 Quarter 6

Refactored Tests To streamline the user-oriented testing of new StratusLab releases, the existing tests have been refactored to provide feedback on particular features and use cases. These refactored tests have been integrated into the Hudson continuous integration system, like their more monolithic predecessors. These provide systematic and rapid feedback for changes in the distribution. Moreover, these tests now use cloud resources to reduce contention for the available, physical testing resources and to avoid interference between different tests running in parallel.

Grid'5000 Initial investigations have been done on using the Grid'5000[5] infrastructure to perform scalability tests of the StratusLab services. Use of the infrastructure has turned out to be more difficult than expected and unfortunately StratusLab services have not yet been deployed within Grid'5000 successfully. Work on this will continue in the next quarter.

[5] https://www.grid5000.fr/mediawiki/index.php/Grid5000:Home

StratusLab operations in a Bioinformatics Lab CNRS IBCP has continued its evaluation of the StratusLab distribution in our bioinformatics context. The release 1.1 is now in production on our local resources. We have deployed the new persistent disk component and have integrated it in our infrastructure. We integrated the PAT (Port Address Translation) implementation with the command line interface. We have begun an evaluation of specific machines with "big memory" responding to specific requirements of bioinformatics tools. We have run bioinformatics instances of 64GB and 96GB with success. We have also evaluated the portability of Intel-defined appliances on AMD CPUs. The "biocompute" appliance, defined on Intel-CPU server, was run on this server without notable differences.

SLCS Authentication and Authorization To fulfill a strong requirement from our scientific community of simple but robust access to our cloud by bioinformaticians, we have deployed in IBCP a authentication solution based on the SLCS framework[6]. This work has been done in collaboration with the Swiss SWITCH team maintaining SLCS software. SLCS is a service to create short live certificates based on Shibboleth authentication and a certificate authority (CA). Certificates are built with user attributes provided by Shibboleth identity provider and should be directly usable on all clouds accepting the CA. No modifications of StratusLab authentication layer are required. We have a working prototype and this is now in production on IBCP's cloud infrastructure.

3.2.3.3 Quarter 7

StratusLab operations in a Bioinformatics Lab CNRS IBCP has continued its evaluation of the StratusLab distribution in a bioinformatics context. The PAT (Port Address Translation) implementation has been integrated with the command line interface and the standard web monitor and is now part of the official release sources of the project. The CNRS IBCP infrastructure has been advertised to some colleagues from the bioinformatics community. Of course, the first ones have been our partner from Institut Pasteur (Paris) with whom we are collaborating for the TOSCANI use case. Two accounts have been created. And they are using our cloud for the evaluation of the ARIA appliance. We have also opened our cloud to colleagues from within IBCP working in the structural biology field. They have requirements of accessing from time to time some computing resources in order to run tools for protein structure modeling. One of these tools has highlighted the need to run software under a license agreement, in that case an academic license. With these first accounts we are beginning our objectives to open our cloud to the bioinformatics community. This action will bring feedback to the project from real usage by non-computing-specialist scientists. But in our mind, this will be better achieved with a dedicated and simple web interface to the VM lifecycle integrating bioinformatics assistants (see related paragraph).

Grid'5000 Deployment Being able to test the StratusLab distribution at scale is critical for its adoption. Resources within the project have limited the level of scal-

[6] http://www.switch.ch/grid/slcs

ability testing that can be done. As a possible solution, a StratusLab deployment inside of the Grid'5000 infrastructure has been investigated. This infrastructure, dedicated to computer science research, allows dynamic deployment of complete systems to the scale of 1500 of machines. An initial deployment was done by hand and validated, requiring only minimal changes in the standard StratusLab configuration. This was then automated, allowing easy redeployments of the system through Hudson. This system will be used as the basis for planned scalability tests in the final quarter.

Use Case Tests Comprehensive testing of the StratusLab features was scattered through a number of different test jobs in Hudson. This made it difficult to determine what features were tested and what was passing/failing. Moreover, many of the tests mixed the installation with client tests making it difficult to ascertain what a "pure" client without elevated access priviledges would experience. A set of distinct use cases covering the expected functionality has been extracted. These are kept in a distinct code repository and run through a job that runs on a machine outside of the test cloud infrastructure. This permits better tracking of the StratusLab distribution's functionality.

3.2.3.4 Quarter 8

Multi-Platform Support The jobs for multi-platform builds have been further refined to provide a complete build chain for Fedora 16, CentOS 6.2, and OpenSuSE 12.1. The Fedora 14 builds are also being maintained until the full set of test jobs are migrated to CentOS 6.2.

Use Case Tests The use case tests continue to be updated with new releases of the software. They are often failing because of changes to the underlying software or because of deployment problems with the services. They remain a good indicator of "releasablility" of the current development version.

Deliverable D2.5 The Deliverable D2.5 summarizes the work done to evaluate the StratusLab software and services over the course of the project. Like the first version of this deliverable, this document compares the functionality provided to initial requirements from several sources.

StratusLab operations in a Bioinformatics Lab CNRS/IBCP has continued its adaptation of the StratusLab distribution in a bioinformatics context. The PAT (Port Address Translation) implementation has been integrated with the new releases of the StratusLab distribution. New computing resources have been added to the IBCP infrastructure, giving it a total of 172 cores, 536 GB of RAM and 25 TB of storage through the persistent disk service. A new instance devoted to "high throughput computing" with 48 CPUs and 120 GB RAM has been defined to satisfy requirements for some bioinformatics tools.

3.2.4 Issues and Corrective Actions

3.2.4.1 Quarter 5

Focus on Use Cases One of the primary recommendations from the reviewers was to refocus the WP2 activities in Y2 to really bring users onto the reference infrastructure by implementing and then publicizing their use cases. The work done for D2.3 was the first step in refocusing of the WP2 activities. Further work will need to be done to define a concrete schedule for the identified use cases and corresponds well with the evolution of the StratusLab distribution.

Delay of MS4 The milestone MS4 concerns the deployment and use of the StratusLab cloud distribution by external sites, that is, sites not associated with one of the partner institutes. There are external sites evaluating the StratusLab distribution (IRES in Strasbourg, FR and RAL in Didcot, UK) but they are not yet running these services in production. Consequently, this milestone has been delayed while we continue to support them during their testing period.

3.2.4.2 Quarter 6

Further Delay of MS4 The MS4 milestone has been further delayed. This is largely because we are waiting for the StratusLab 1.2 release that contains a large number of new features. That release has been delayed by serious air conditioning problems at LAL in this quarter and because of the need to streamline the build and test procedures. We expect production use of the 1.2 release by external sites in the next quarter.

3.2.4.3 Quarter 7

Signficant progress was made on getting specific applications running on StratusLab cloud infrastructures; however, these successes have not been disseminated widely through a press release or public presentation. In Q8, this progress will be presented at the EGI Community Forum, but a formal press release and wider publicity needs to be planned. This will be a priority for Q8.

As seen in MS4 there are a good number of sites that have deployed StratusLab both inside and outside of the project. However, there is a need to push more sites to deploy StratusLab and to provide production cloud services to their users. The work by SixSq for commercial development will help, but must be complemented by specific actions within the academic community.

3.2.4.4 Quarter 8

No particular issues have arisen in this quarter. However, a point will have to be made to disseminate the successful use of the StratusLab cloud infrastructures for a wide range of applications during the v2.0 dissemination activities.

3.3 WP3: Dissemination

Work Package 3 coordinates the project's activities in dissemination, collaboration, exploitation and sustainability. Its objectives are to disseminate results of the project to resource providers, end-users, and the general public; identify project contributions to standards bodies and standardization efforts; and coordinate interactions with related projects, developing Memoranda of Understanding between projects where appropriate.

3.3.1 Summary

3.3.1.1 Quarter 5

Dissemination efforts have continued with the release of a new updated website, press releases and articles in relevant publications relating to the release of StratusLab version 1.0 as well as the StratusLab Marketplace. Two hands-on training workshops were held along with a large number of talks, presentations and panel discussions where the StratusLab project and software was described.

Two new Memoranda of Understanding have been signed this quarter, with CYFRONET and VENUS-C. A draft MoU with IGE has also been drawn up, and progress continues to be made in line with the existing MoUs. Collaborations with other projects have also continued.

The project's exploitation and Sustainability plan has now been developed and progress towards these goals has begun, focussing on the work of the project's commercial partners SixSq Sàrl and Telefónica I+D, as well as initiating contact with other commercial companies.

3.3.1.2 Quarter 6

Dissemination efforts have continued this quarter and though press coverage and website visits are down on Q5, the number of training events increased with three very well-attended hands-on tutorial sessions and a 89 people trained in the StratusLab software this quarter. A large number of talks and presentations featuring StratusLab were also given.

The project has put in place a new announcement mailing list and a user discussion forum which it is hoped will form the hub of a community support mechanism.

The project had great success at the EGI Technical Forum in September, with a booth and a number of presentations and participation in technical sessions. In addition the StratusLab poster "Virtualization of Bioinformatics Applications on Cloud Infrastructures" won the runner-up prize in the Best Poster competition.

A Memorandum of Understanding was signed with the IGE project, and a draft MoU has been drawn up with ERINA+ which will give StratusLab access to the impact assessment process of that project.

Work on Exploitation and Sustainability has increased, with participation in the DCI Sustainability meeting in October and continuing work within the project to identify avenues for sustainability.

3.3.1.3 Quarter 7

Dissemination efforts have continued this quarter with activities centered around the release of versions 1.2 and 1.3. The release announcements were widely circulated via the project mailing lists and were picked up by some technical press outlets. A number of talks and presentations featuring StratusLab were also given by project partners.

The website continues to be the main avenue for dissemination to the users, and the documentation section has been extensively updated during this quarter. While site visits are down again in Q6, this is to be expected due to the holiday period falling within Q7.

Significant progress has been made with collaborating projects, in particular with EGI who have deployed their own instance of the StratusLab Marketplace as part of their Federated Clouds Task Force activities, and new EGI-endorsed UMD virtual appliances available in the StratusLab Marketplace. Collaborations with other projects have also continued.

Exploitation and Sustainability efforts have increased, with participation in the EGI Sustainability Workshop in January, and progress on a number of commercial exploitation initiatives.

3.3.1.4 Quarter 8

Dissemination efforts have continued this quarter even as effort has turned towards sustainability and other activities in preparation for the end of the project. The release announcement for version 1.4 was widely circulated and there has been some publicity surrounding commercial exploitation efforts of the partners.

Website visits are up over the previous quarter and use of social media and mailing lists has also increased. The use of the StratusLab user forum, however, is still low.

Most formal collaborations will draw to a close at the end of the project but informal collaborations that make use of the results of the StratusLab project are expected to continue, driven by individual partners, or by the open-source StratusLab community.

A final plan for exploitation and sustainability has been put in place which allows for the continued development and maintenance of the StratusLab Distribution and other outputs of the project beyond the project lifetime.

3.3.2 Task 3.1: Dissemination

3.3.2.1 Quarter 5

Release Dissemination The release dissemination plan was updated for the major release 1.0. EGI.eu published a short article 'The Grid on a Cloud'[7] on the project and new release, and International Science Grid This Week (iSGTW) carried an annoucement of the release.[8] CNRS produced a press release in French

[7] http://www.egi.eu/about/news/news_0068_StratusLabSoftwareV1.html

[8] http://www.isgtw.org/announcement/stratuslab-releases-open-source-cloud-solution-ready-grid

'Cloud computing et grilles informatiques : lancement de Stratuslab 1.0' was picked up and reprinted by a number of french technical media outlets,[9] with an article appearing in the French technology outlet L'Atelier BNP Paribas[10]

StratusLab v1.0 was featured on the SixSq website[11] in conjunction with a commercial support offering.

Further afield, the news of the release was carried by HPC in the Cloud[12] and HPCwire. [13]

Release 1.1 is due in early September and release dissemination will coincide with the EGI Technical Forum 2011.

Media & Publications StratusLab publications in the media in Q5 have been associated with the release of the production-ready StratusLab v1.0. As mentioned above, both the English and French language press releases were widely picked up and reprinted, while articles about the release were published by EGI.eu and L'Atelier BNP Paribas.

The release of the StratusLab Marketplace was highlighted and publicized as a news release by the Trinity College Dublin communications office.[14] which was picked up by the Irish technical online media outlet TechCentral.ie[15], as well as being covered in a number of Irish Education and Research-focussed websites.

Many of these media stories were also widely publicised by project members and related projects via Twitter.

A comprehensive list of media and publications is maintained on the project website.[16]

Website The redesigned website[17] went live to coincide with the release of version 1.0.

Figure 3.2 shows the number of visits to the website. The number for Q5 (5,472 visits) is up almost 20% from Q4 (4,579 visits).

The StratusLab Twitter feed now has 74 followers.

EGI Technical Forum 2011 The project booked an exhibition booth at EGI Technical Forum 2011, to be held in Lyon, France on 19–23 September 2011. The booth will feature posters, demonstrations of the software. The project is also planning dissemination material including t-shirts, brochures, and pens.

Project members are chairing the Virtualization & Cloud Computing and Agile

[9] http://www.in2p3.fr/recherche/nouvelles_scientifiques/2011/9_stratuslab.htm

[10] http://www.atelier.net/articles/cloud-aider-organismes-de-recherche-gerer-leurs-pics-de-charge

[11] http://www.sixsq.com/front-page/stratuslab

[12] http://www.hpcinthecloud.com/hpccloud/2011-07-05/stratuslab_releases_open-source_cloud_solution_designed_for_the_grid.html

[13] http://www.hpcwire.com/hpcwire/2011-07-05/stratuslab_releases_open-source_cloud_solution_designed_for_the_grid.html

[14] http://www.tcd.ie/Communications/news/news.php?headerID=1976&vs_date=2011-8-1

[15] http://www.techcentral.ie/article.aspx?id=17220

[16] http://www.stratuslab.eu/doku.php/press

[17] http://www.stratuslab.eu

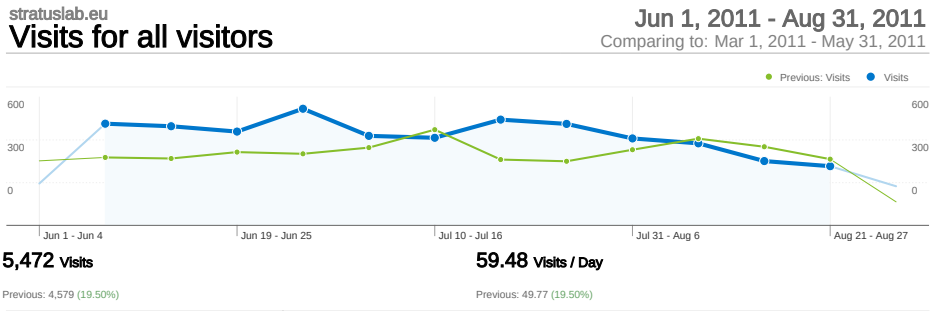

Figure 3.2: *Visits for Q5.*

Methodologies sessions at the Technical Forum. The project will present its status and developments in the Virtualization & Cloud Computing session. StratusLab will give two 90 minute training sessions at the forum, organized by WP2.

French annual Bioinformatics Conference CNRS IBCP has presented a poster at the French annual Bioinformatics Conference JOBIM 2011, hosted by Institut Pasteur, Paris, from 28 June to 1st of July 2011. This poster, titled "Virtualisation of Bioinformatics Applications on Cloud Infrastructure", presented the virtual bioinformatics appliances that have been built by the partner CNRS and the benefits to bioinformatics scientists and engineers of using the cloud service from StratusLab. Contacts with bioinformatics researchers have been made during the discussions around the poster.

Talks Two hands-on training tutorials were organised during Q5, as well as a number of talks describing the project. A brief list is given in Table 3.1. Details and links, where available, are given on the project website[18].

The project plans to submit a proposal to the EU Innovation Convention to be held on 5–6 December 2011, Brussels

3.3.2.2 Quarter 6

Demos The project is developing a number of demonstrations of the StratusLab software, use cases and features. These include a live interactive user-focused demo as well as several video demos of the functionality of the StratusLab software.

A proposal was submitted for a demo at the EU Innovation Convention to be held on 5–6 December 2011 in Brussels. The proposal was not accepted, but work on use cases to demonstrate continues and it is hoped that these will be used at future events.

Media & Publications StratusLab was featured in a piece in International Science Grid This Week (iSGTW) on 11 November entitled "European initiatives strengthen cooperation on cloud computing" which described how DCI projects

[18] http://stratuslab.eu/doku.php/presentations

Table 3.1: *Talks (Q5)*

Title / Event	Date
Grid-Ireland Operations Centre / TCD presentation featuring StratusLab at Lero Cloud Computing Workshop Dublin, Ireland	2011-06-03
"What's the deal with Agile Contracts" Panel discussion(C. Loomis) at XP2011, Madrid, Spain	2011-06-11
"StratusLab : Les application scientifiques sur le cloud" (C. Loomis) at "Calcul intensif pour la biologie" in Lille, France	2011-06-14
"Inforum, Le Cloud en Marche (working cloud)" Stand, posters and demos at Palexpo, Geneva, Switzerland	2011-06-15
"Introduction au Cloud Computing et le projet StratusLab" (C. Loomis), seminar in Marrakech, Morocco	2011-06-17
"Building Clouds with OpenNebula 2.2 and StratusLab", tutorial at CONTRAIL Summer School 2011, Hyères-les-Palmiers, France	2011-06-27
"Challenges in Federated and Hybrid Cloud Computing", keynote at the 2011 International Conference on High Performance Computing & Simulation, Istambul, Turkey	2011-07-06
"OpenNebula and Stratuslab toolkits for virtualization", talk at CloudCamp Valencia 2011, Valencia, Spain	2011-07-13
StratusLab training day Joint European DCI Summer School, Budapest, Hungary	2011-07-15
"Challenges in Federated and Hybrid Cloud Computing", keynote at the 2011 International Conference on Parallel Computing, Ghent, Belgium	2011-08-31

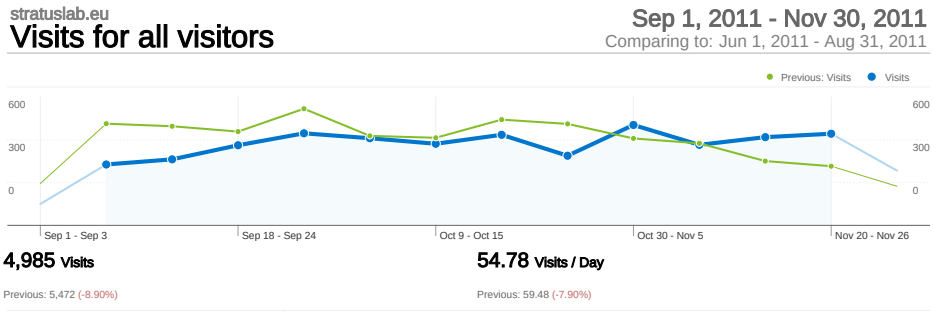

Figure 3.3: *Visits for Q6.*

are cooperating to enhance European Cloud infrastructures for research.

It was intended to produce some material for the popular press, but this has not yet been done and will happen instead in the next quarter.

Website, Mailing Lists and Fora Figure 3.3 shows the number of visits to the website. The number for Q6 (4,985 visits) is down slightly from Q5 (5,472 visits).

The StratusLab Twitter feed now has 87 followers.

The announcements mailing list, which allows interested members of the public to keep up-to-date with news from the StratusLab project such as new releases, now has 72 members.

In October an additional "user announcements" mailing list was created, specifically for users of the StratusLab reference infrastructure. This is an announcement only mailing list used to inform users of upcoming service outages, upgrades and other important news. The list has 46 members, all of whom are users of the reference infrastructure.

A new public user forum was created and went live on 16 November, and has been running on a trial basis so far. The aim of the forum is to provide a contact point for StratusLab users and to build a community support mechanism. The idea of the forum appears to be feasible and the group already has 12 members, and has seen seven messages posted so far in two topics. It will thus be officially launched in December 2011 and disseminated more widely to StratusLab users and related communities.

EGI Technical Forum 2011 StratusLab had a very visible presence at the EGI Technical Forum which was held in Lyon, France from 19–23 September. The project had a booth displaying posters on Agile Methodology, Grid-Cloud Integration, Hadoop on StratusLab, The StratusLab Marketplace, StratusLab Reference Infrastructure, SlipStream and Cloud Storage. StratusLab-branded pens and post-it notes were printed and distributed from the booth. The project was also well represented in the technical programme of the conference with partner TCD's David O'Callaghan co-chairing sessions in the Cloud and Virtualisation track.

Project members also gave four presentations at the conference covering topics such as cloud deployment, Agile Development, Bioinformatics on the cloud and

Standards in StratusLab.

Christophe Blanchet of partner CNRS/IBCP won the runner-up prize in the poster competition for his poster "Virtualization of Bioinformatics Applications on Cloud Infrastructures" and a successful hands-on tutorial session was also held.

Talks and Events Q6 saw an increase in the number of training events, with three hands-on tutorials held. The first at the EGI Technical Forum in September attracted about 25 participants. StratusLab participated in the 3rd School on Advanced Computing and GRID Technologies for Research (ACGRID3) in Hanoi, Vietnam in November, where 35 students were trained in using StratusLab over three days. Finally a two-day training event dedicated to StratusLab was held in November at the Linear Accelerator Laboratory in Orsay, France. This was attended by 25 people and the event was also streamed live on the internet with three participants following online. In total 89 people were trained in using StratusLab this quarter.

A large number of dissemination talks mentioning StratusLab were also given. A full list of talks and events is given in Table 3.2.

3.3.2.3 Quarter 7

Release Dissemination There were two incremental releases of the StratusLab distribution during this quarter, version 1.2 and 1.3. A release dissemination plan was produced for each of these releases. Notification of the release was disseminated via the website and the project's announcement email list. The release announcements for both releases were also sent to various media outlets and were carried by International Science Grid This Week (iSGTW) and HPC In the Cloud.

Media & Publications StratusLab participated in the Cloudscape IV event in February. News of the project's participation was posted on the website and was also disseminated to press contacts. The story was picked up by HPC In the Cloud.

Partner project the Initiative for Globus in Europe (IGE) ran a news item about the availability of Globus Appliances from the StratusLab Marketplace.

The release announcements and other media stories, events and general project news were also disseminated via Twitter, with several items being retweeted by related projects.

Website, Mailing Lists and Fora Figure 3.4 shows the number of visits to the website. The number for Q7 (4,442 visits) is down slightly again from Q6 (4,985 visits).

The StratusLab Twitter feed now has 101 followers.

The announcements mailing list, which allows interested members of the public to keep up-to-date with news from the StratusLab project such as new releases, has 72 members.

The user announcements mailing list for users of the StratusLab reference infrastructure has 46 members, all of whom are users of the reference infrastructure.

The public user forum has 13 members, although there have been no posts on the forum during this period.

Table 3.2: *Talks (Q6)*

Title / Event	Date
"Challenges in Federated and Hybrid Cloud Computing" at the 2011 International Conference on Parallel Computing, Ghent, Belgium	2011-08-30 – 2011-09-02
"Challenges in Federated and Hybrid Cloud Computing" at Cloud Day 2011, Stockholm, Sweden	2011-09-14
"StratusLab: status update and elastic deployment of services" at EGI Technical Forum, Lyon, France	2011-09-20
"StratusLab, Bioinformatics and the cloud" oral presentation at the EGI Technical Forum	2011-09-20
"Agile in StratusLab" at EGI Technical Forum, Lyon, France	2011-09-20
"StratusLab project: Standards, Interoperability and Asset Exploitation" at SIENA Initiative Workshop, EGI Technical Forum 2011, Lyon, France	2011-09-21
"Next Generation IaaS Cloud Computing with OpenNebula 3.0" at ISOD Workshop, OGF33, Lyon, France	2011-09-21
"StratusLab Tutorial" at EGI Technical Forum, Lyon, France	2011-09-22
"Virtualization of Bioinformatics Applications on Cloud Infrastructures" poster presentation and winner of runner-up prize in Poster Competition, EGI Technical Forum, Lyon, France	2011-09-19 – 2011-09-23
"StratusLab Booth" at EGI Technical Forum, Lyon, France	2011-09-19 – 2011-09-23
"Building Private Clouds for HPC with OpenNebula: Reference Deployments & Lessons Learned" at ISC Cloud 2011, Mannheim, Germany	2011-09-27
"StratusLab: Cloud Federation" at Federating IaaS Cloud for UK Research Workshop during the UK e-Science All Hands Meeting, York, UK	2011-09-28
"StratusLab Marketplace" at Atelier France Grilles, Lyon, France	2011-10-21
"OpenNebula Interoperability and Portability" at 5th International DMTF Academic Alliance Workshop on Systems and Virtualization Management: Standards and the Cloud, Paris, France	2011-10-24
"StratusLab Tutorial" at 3rd School on Advanced Computing and GRID Technologies for Research (ACGRID3) Hanoi, Vietnam	2011-11-01 – 2011-11-03
"StratusLab Tutorial" Orsay, France	2011-11-17 – 2011-11-18
"Virtualisation de serveurs l'aide d'un logiciel de cloud" JRES 2011 in Toulouse, France	2011-11-23

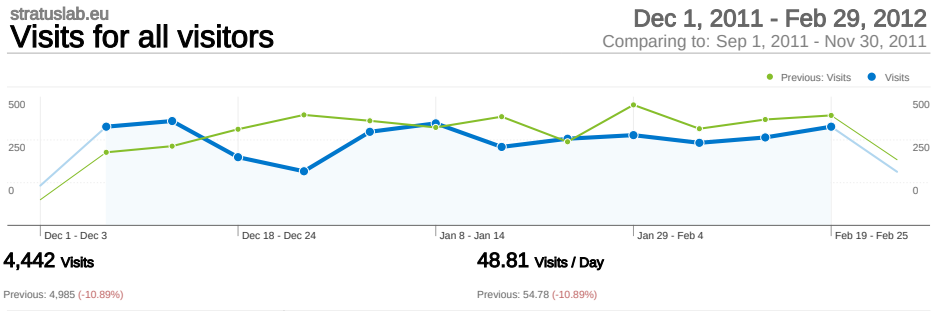

Figure 3.4: *Visits for Q7.*

EGI Community Forum 2012 The project has booked an exhibition booth at the EGI Community Forum 2012, to be held in Munich, Germany 26-30 March 2012. The booth will be the centre of dissemination activities with a number of posters highlighting successes of the project. Promotional items such as t-shirts, brochures, pens and post-it notes will also be distributed. Video demos will be shown on a large screen at the booth, and a live demo will be available.

Project members will be involved in several of the conference sessions on Cloud computing. A StratusLab training workshop is also scheduled.

Talks A brief list of talks featuring StratusLab given by project members is given in Table 3.3.

3.3.2.4 Quarter 8

Release Dissemination StratusLab released version 1.4 during this quarter. A release dissemination plan was produced for this release. Notification of the release was disseminated via the website and the project's announcement email list. The release announcements for both releases were also sent to various media outlets and were carried by International Science Grid This Week (iSGTW) and HPC In the Cloud.

Media & Publications A press release was written about the release of the DS-Cloud Ready Pack by partner SixSq in collaboration with Swiss IT Solutions Provider Darest. This was widely disseminated and picked up by several online media outlets.

The release announcements and other media stories, events and general project news were also disseminated via Twitter, with several items being retweeted by related projects.

A major dissemination effort is expected just beyond the end of the project to coincide with the final release of StratusLab v2.0.

Website, Mailing Lists and Fora Figure 3.5 shows the number of visits to the website. The number for Q8 (5,372 visits) is up from Q7 (4,442 visits).

The StratusLab Twitter feed now has 101 followers.

The announcements mailing list, which allows interested members of the pub-

Table 3.3: *Talks (Q7)*

Title / Event	Date
"Experience from the provision of IaaS cloud services using the StratusLab distribution" at eScience 2011, Stockholm, Sweden	2011-12-07
"What benefits of cloud for Bioinformatics - a user perspective" at eScience 2011, Stockholm, Sweden	2011-12-07
"What can grid and cloud computing do for you?" Seminar in School of Computer Science and Statistics, TCD, Dublin, Ireland	2012-01-27
"Grid and Cloud e-Infrastructures" at meeting with Intel Academic Research Programme representatives, Dublin, Ireland	2012-01-30
"SlipStream - Automated provisioning and continuous deployment in the cloud" at CERN, Geneva, Switzerland	2012-02-01
"OpenNebula Project" featuring StratusLab at FOSDEM, Brussels, Belgium	2012-02-04
"Open-source and Standards - Unleashing the Potential for Innovation of Cloud Computing" at CloudScape IV, Brussels, Belgium	2012-02-23
"StratusLab project status" presentation at round-table session at CloudScape IV, Brussels, Belgium	2012-02-23

lic to keep up-to-date with news from the StratusLab project such as new releases, has 73 members.

The user announcements mailing list for users of the StratusLab reference infrastructure has 47 members, all of whom are users of the reference infrastructure.

The public user forum has 16 members, although there has again been relatively little use of this forum, despite attempts to disseminate the url more widely.

Booth at EGI Community Forum The project prepared a booth for the EGI Community Forum event in Munich, Germany. Four posters were prepared covering the StratusLab service, applications using StratusLab, the trusted appliance "ecosystem" and our use of agile software development methodologies. A large plasma screen provided a continuous loop of videos from the project. In addition, an updated project brochure, a take-home flyer of the posters, and copies of the StratusLab overview paper were provided. The entire run of 100 StratusLab t-shirts were given away during the event.

Talks A brief list of talks featuring StratusLab given by project members is given in Table 3.4.

Table 3.4: *Talks (Q8)*

Title / Event	Date
"Cloud, Ready for Bioinformatics?" at the EGI Community Forum, Munich, Germany	2012-03-28
"Supporting grid-enabled GPU workloads using rCUDA and StratusLab" at the EGI Community Forum, Munich, Germany	2012-03-28
"StratusLab: Use cases, features and sustainability" at the EGI Community Forum, Munich, Germany	2012-03-29
"StratusLab Training" at the EGI Community Forum, Munich, Germany	2012-03-29
"Operating a Public Cloud with StratusLab" at the EGI Community Forum, Munich, Germany	2012-03-29
"Public vs Private Cloud Usage Costs: The StratusLab Case". 2nd International Workshop on Cloud Computing Platforms (CloudCP) - In conjunction with EuroSys 2012, Bern, Switzerland	2012-04-10
"The StratusLab distribution: Use-cases and support for scientific applications" at Earth & Space Science Information session 2.10, EGU General Assembly 2012, Vienna, Austria (Invited Talk)	2012-04-26
"An open source cloud solution based on KVM" Pierre Perdeams (IBM) / Marc-Elian Begin (SixSq) at the IBM Technical University, Birmingham, UK	2012-05-10

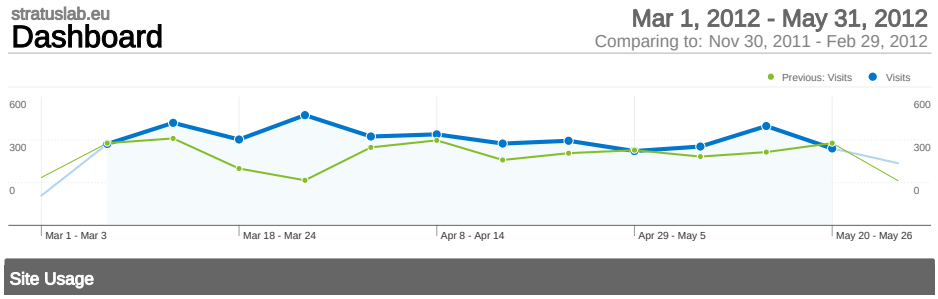

Figure 3.5: *Visits for Q8.*

3.3.3 Task 3.2: Collaboration with Standards Bodies and Related Projects

3.3.3.1 Quarter 5

Memoranda of Understanding StratusLab has signed new Memoranda of Understanding with Cyfronet, VENUS-C.

CYFRONET StratusLab has signed a new MoU with the Academic Computer Centre CYFRONET of the University of Science and Technology in Krakow, Poland. Under the terms of the MoU, the two parties will share knowledge related to cloud and virtual technologies and StratusLab will provide a software platform and support to CYFRONET for the Polish Grid infrastructure, while CYFRONET will provide requirements to the project.

VENUS-C StratusLab has signed a new MoU with the VENUS-C project. VENUS-C will deploy OpenNebula and provide feedback and requirements to StratusLab based on their experiences. In addition to this the two projects will aim to make their software inter-operable on the interface level and to share experiences in the area of accounting. StratusLab will also investigate the use of the VENUS-C storage solutions.

EGI Under its MoU with EGI, StratusLab has contributed to the EGI prioritised list of standards by identifying the standards which the project implements or intends to implement. StratusLab is also continuing to provide input on the EGI UMD roadmap and is represented in the EGI Security Coordination Group, and the EGI training workgroup.

EMI Discussions have continued with EMI where StratusLab is in close collaboration with the EMI Virtualization group, providing comments on the virtualization-enabled architecture proposed by EMI. Other areas of collaboration include possible integration on the security level, such as adoption of Argus authorization framework, support for SAML, etc. It is planned to discuss these areas further with EMI project members at the EGI Technical Forum in Lyon in September.

Contrail StratusLab has been in contact with Contrail and hope to work together to share code and modules and push for standards for authorisation and accounting.

IGE A draft MoU has been drawn up between StratusLab and IGE which allows for the deployment of Globus services and tools on the StratusLab platform and creation of Globus Appliances for the StratusLab Marketplace. IGE will provide feedback on their use of the StratusLab tools as well as requirements for future developments. In addition the two projects will collaborate to deliver common training events.

SIENA In July StratusLab provided an update on the project after one year in development and operation. StratusLab also continues to provide SIENNA with input, for example, on the recent SIENNA gap analysis document. Project partner Vangelis Floros is participating in SIENNA phone calls as member of the REB, and in the SIENNA session at the EGI Technical Forum in September.

European Cloud Computing Infrastructure Project representatives attended the European Cloud Computing Strategic Plan workshop at ESA/ESRIN from 28–29 June. This workshop aimed to develop an initial vision and direction for a European Industrial Strategy for a Scientific Cloud Computing Infrastructure to be implemented by 2020, and included research and academic stakeholders as well as commercial cloud providers. StratusLab will continue to engage with this group and will participate in a follow-up Supply-Side (cloud provider) meeting in September.

3.3.3.2 Quarter 6

Memoranda of Understanding

VENUS-C As part of the MoU signed between StratusLab and VENUS-C project, VENUS-C is evaluating OpenNebula with support from StratusLab, and StratusLab is evaluating storage and accounting solutions provided by VENUS-C.

EGI Collaboration with EGI continues in line with the MoU established with that project. An EGI GGUS support team to support sites running StratusLab was established in September which the projects hope will encourage more sites to run the StratusLab software.

StratusLab provided training and technical talks, as well as co-chairing a session at the EGI Technical Forum in September 2011 and the project has submitted a number of session abstracts (including another tutorial session) for the EGI Community Forum 2012. StratusLab has also committed to participate in technical integration sessions organised by EGI at the Community Forum.

Talks are in progress with EGI on the possibility of EGI using or running their own instance of the StratusLab Marketplace. This may provide a possible avenue for sustainability of the Marketplace.

StratusLab participated in a meeting on Sustainability of DCI projects organised by EGI in October. StratusLab is also providing input to the UMD roadmap, in particular EGI have asked for input in relation to messaging/notification integra-

tion.

EMI The technical collaboration with the EMI project has continued in Q6. StratusLab provided input to the EMI presentation on Cloud computing at the EGI Technical Forum in Lyon, France in September. In addition, the two projects worked together to prepare Torque images which were requested by EMI and have now been made available in the StratusLab Marketplace. StratusLab also provided support to EMI on the utilization of these VM images.

IGE StratusLab signed an MoU with the IGE project at the EGI Technical Conference in Lyon, France in September. The MoU allows for the deployment of Globus services and tools on the StratusLab platform and creation of Globus Appliances for the StratusLab Marketplace. IGE will provide feedback on their use of the StratusLab tools as well as requirements for future developments. In addition the two projects will collaborate to deliver common training events.

SIENA StratusLab participated in the SIENA workshop at the EGI Technical Forum in Lyon, France in September, where Vangelis Floros of partner GRNET gave a presentation describing StratusLab's use of Standards and Interoperability.

The SIENA project also coordinated input from EU projects to the US National Institute of Standards and Technology (NIST) Cloud Computing Technology Roadmap, and StratusLab provided input in October to SIENA on how the projects solutions map to the NIST use cases and also participated in a number of phone conference meetings on this topic.

ERINA+ A draft MoU has been prepared between StratusLab and the ERINA+ project (which was signed in December). The collaboration will give StratusLab access to the impact assessment tools and process developed by ERINA+, in order to define a common methodology for the evaluation of impact of e-Infrastructures investments. It will allow StratusLab to measure the impact of the project's work in a methodologically sound way in collaboration with experts.

ERINA+ will provide a self-assessment online tool which is expected to be available from February 2012.

Mantychore The Mantychore project has expressed continued interest in liaising with StratusLab and Mantychore grid/cloud integration use cases will employ the StratusLab distribution.

3.3.3.3 Quarter 7

VENUS-C The project has continued its collaboration with VENUS-C during this quarter. Feedback was received from Venus-C users about requirements and user experience of the IaaS approach and this feedback was used to direct further development. In addition, the two projects have continued to share experiences relating to installation, performance and use of OpenNebula.

Some investigation has been done on the storage solution provided by VENUS-C and its relevance to StratusLab. Although this could probably be used to wrap the StratusLab storage services, the utlity of this is questionable especially in view

of the need to improve the performance of the storage system in general.

EGI Collaboration with EGI continues in line with the MoU established between the two projects.

EGI have deployed their own instance of the StratusLab Marketplace, and evaluation of its use is ongoing. The StratusLab project undertook some development of the Marketplace in order to allow the EGI instance[19] to use its own branding, and this has been completed.

Partner TCD is participating in the EGI Federated Clouds Task Force as a resource provider with their cloud site running StratusLab software. TCD are also working with the Task Force to enable Nagios monitoring of TCD's StratusLab installation through OCCI v1.1. Minor modifications of the Nagios probe to make use of the StratusLab Marketplace have been needed.

StratusLab is planning to have a strong presence at the EGI Community Forum which will take place in March 2012. The project will take part in the Cloud and Virtualisation track at the conference and has submitted a training session and several presentations for the event.

Project partners participated in the EGI Sustainability Workshop which was held in Amsterdam, Netherlands, from 24-26 January 2012.

StratusLab has also provided input to EGI on joint dissemination activities.

IGE Work has continued with the Initiative for Globus in Europe (IGE) project, with IGE contributing a number of fully tested and endorsed Globus Toolkit 5.0.4 appliances to the StratusLab Marketplace.

SIENA StratusLab had a strong presence at the SIENA CloudScape IV event. Project members gave two talks and participated in a number of round-table discussion sessions.

The project also provided input this quarter to the SIENA Roadmap on Distributed Computing Infrastructure for e-Science and Beyond in Europe.

ERINA+ An MoU was signed between StratusLab and ERINA+ in December 2011. As part of this MoU, StratusLab has committed to participating in the ERINA+ socio-economic impact assessment. ERINA+ has provided the project with access to their online self-assessment tool and StratusLab is currently in the process of collecting the necessary information to complete the self-assessment. This will allow StratusLab to measure the impact of the project's work in a methodologically sound way in collaboration with experts.

Commercial collaborations Partner SixSq has been involved in a collaboration with IBM and Swiss IT solutions provider Darest to develop a commercial turnkey Infrastructure as a Service (IaaS) cloud solution. The product, which will be distributed by Darest and which runs on IBM hardware, uses StratusLab 1.4 as its core, and is aimed at SMEs looking for a private cloud solution. The product will be announced in March.

[19] Available at http://marketplace.egi.eu/.

3.3.3.4 Quarter 8

ERINA+ StratusLab has attempted complete the ERINA+ Socio-Economic Impact Self-Assessment of e-infrastructure Projects using the online self-assessment tool provided by ERINA+. Initial results of the assessment are relatively favourable, but updates to the tool and some technical issues have impacted our ability to complete the assessment process.

StratusLab also participated in the ERINA+ session at the EGI Community Forum in March.

VENUS-C Support to Develop OVF4ONE This software (http://grids15.eng.it/svn/main/ovf4one) is a Java implementation of OCCI according to GFD.184 using OVF messages and OpenNebula as backend. ovf4one translates RESTful calls to the OCCI methods into OCA RESTful calls, and the OVF XML message is translated into OpenNebula VM templates. ovf4one is part of the OpenNebula software ecosystem[20].

VENUS-C Support, Feedback and Experience Sharing Members of VENUS-C have received support about OpenNebula (contextualization process, use of public Cloud APIs, Cloud federation..). VENUS-C has provided feedback to StratusLab about OpenNebula, including feature requests and bugs, and the StratusLab IaaS model. Both projects have shared experiences installing and deploying OpenNebula.

Support and Integration for Spanish NGI During this period StratusLab has assisted the Spanish-Portuguese NGI to set up several cloud sites to build a federated cloud infrastructure. Along with StratusLab components, such as OpenNebula and the Marketplace, other IaaS platforms are used (e.g. OpenStack). The UCM team also helped with the integration of the authentication system with the LDAP services of the sites by developing specific multi-site authentication drivers.

EGI Collaboration with EGI has continued, including delivery of a StratusLab tutorial at the EGI Community Forum 2012 and participation in the EGI Federated Cloud Task Force, where the StratusLab Marketplace is in use.

The project is also providing input to EGI for a proposed session on business models at the EGI Technical Forum which will take place in September 2012 in Prague, Czech Republic. StratusLab will provide examples of potential and actual business models for organisations using the StratusLab Distribution and other outputs of the project. These examples will draw upon the real commercial products and use cases which have been developed by the StratusLab project partners.

SIENA The SIENA project StratusLab submitted a response to the SIENA *call for actions* outlining how the project has and will continue to contribute to a range of initiatives, including:

- collection of deployoment scenarios

[20] http://www.opennebula.org/software:ecosystem:ovf4one

- collaboration and standards
- contributing to the European Digital Market
- cloud federation
- developing and identifying business models
- sustainability and reuse of project outputs beyond the project
- inventory of grid and cloud standards for research

3.3.4 Task 3.3: Development of Exploitation and Sustainability Plan

3.3.4.1 Quarter 5

The initial exploitation and sustainability plan was delivered in PM10. Since then work in this area has begun following the plan set out in that document.

During Q5 SixSq Sàrl released the first version of its SlipStream product for automated creation virtualised system testbeds with fully integrated support for StratusLab version 1.0. This provides StratusLab with a new use case and potential new user community. Partner Telefónica I+D has also continued to develop the Claudia service manager within the StratusLab project, as part of the Telefónica I+D Global Cloud Initiative.

StratusLab partners have made contact with local business consultancy organisations as part of an effort to develop a more comprehensive business plan.

Initial contacts have also been made with some commercial companies interested in the StratusLab Distribution including Axceleon, a cloud computing consultancy firm.

3.3.4.2 Quarter 6

StratusLab participated in a meeting on Sustainability of DCI projects organised by EGI in October where the project presented its existing sustainability plan and the assets produced by the project.

Work continued at the StratusLab face-to-face meeting in Dublin from 30 November – 2 December on identifying the sustainability path for each of the project's assets. The discussions at this meeting will form the basis of the final Exploitation and Sustainability plan.

In addition to the work above, the commercial partners within the project have continued to exploit the outputs of the project. SixSq now provides a version of its SlipStream product which runs on StratusLab, partner Telefónica I+D has developed a use case for StratusLab based on cloud deployment of an e-commerce site.

Partner TCD is in discussions with the School of Business in Trinity College Dublin about the possibility of using StratusLab as a use case for business students who would create a business plan for StratusLab as part of their final-year projects.

3.3.4.3 Quarter 7

StratusLab participated in the EGI Sustainability Workshop which was held in Amsterdam, Netherlands, from 24-26 January 2012. The sustainability path developed during the meeting in Dublin from 30 November to 2 December provided input to this meeting.

Partner Telefónica I+D is developing the commercial use case in StratusLab to show the advantages of the StratusLab multi-tier service management, KPIs driven scalability and load balancing for an e-business application. By using this proof of concept, TID can show the advantages of using these advanced features that StratusLab provides.

As part of the StratusLab commercial exploitation strategy, partner SixSq has been involved in a collaboration with IBM and Swiss IT solutions provider Darest to develop a commercial turnkey Infrastructure as a Service (IaaS) cloud solution. The StratusLab distribution is one of the key components of this solution.

The product, distributed by Darest and running on IBM hardware, will be announced in March and is specifically tailored for at SMEs looking for a private cloud solution. This brings the StratusLab software to a new market beyond the mainly research-oriented current user base.

An instance of the StratusLab Marketplace will also be made available as a managed service for users of the new product.

3.3.4.4 Quarter 8

Initial Sustainability Actions As part of the sustainability plan currently running services have to be moved to locations that provide the possibility of long-term support. Similarly documentation and materials need to be archived in public locations. Some of these actions are already underway.

- The code repositories have been moved from a dedicated server to GitHub, allowing current members to access the repositories after the end of the project and other people in the community to contribute.
- All project deliverables and milestones have been unloaded to an open access repository (HAL). This will ensure that these documents are visible and available for the long-term.

Other actions and long-term development of a community around StratusLab will occur after the project's close.

The project has written its "Exploitation and Sustainability Final Plan" (D3.5) which outlines in more detail the exploitation outcomes and plans for sustainability beyond the project life time. D3.5 is due concurrently with this report.

3.3.5 Issues and Corrective Actions

3.3.5.1 Quarter 5

The project review conducted in July 2011 highlighted some issues and included a number of recommendations related to this Work Package. These, along with

the corrective actions to be taken are described in the Review Recommendations section of this document. WP3 will attempt to become more marketing-driven in its approach to dissemination and to focus on specific user-communities, and develop demos and dissemination materials aimed at these communities.

3.3.5.2 Quarter 6

The project review in July 2011 called for an increase in dissemination to a wider audience, including the general public. Work has begun on writing articles for the popular science media, however, no such material has been published to date.

Indeed the overall press coverage for this quarter was quite low. This is partly due to the fact that only a single mainly bug-fixing release was made. The upcoming release of version 1.2 will be accompanied by a press release and it is expected that there will be more press coverage in Q7.

3.3.5.3 Quarter 7

The overall activity both within WP3, and in terms of visits to the websites, mentions in press, etc. is down this quarter when compared to the previous quarter. This is largely due, however, to the holiday period falling in this quarter and the trend can be expected to reverse in Q8.

A more concerted effort will be made in Q8 to drive traffic to the website through dissemination, news stories in the press, etc. In particular, the user forums received very little traffic and no posts during this quarter and these must be more widely disseminated in order to transform them into a useful community resource.

3.3.5.4 Quarter 8

It was hoped that the ERINA+ Socio-Economic Impact Self-Assessment could be fully completed before the end of the project, however, as the self-assessment tool is still in development, the results of the assessment are not yet available to be included in the StratusLab final deliverables.

Usage of the StratusLab User Forum is still low despite efforts to increase the visibility of the forum. As the StratusLab Distribution will become an open-source project going forward the collaboration will continue to publicise the user forum and encourage its use by the user and developer community.

It is worth mentioning that a final dissemination effort is expected with the imminent release of StratusLab v2.0, although the results of this are not reported here.

3.4 WP4: Software Integration and Distribution

This activity integrates and supports the StratusLab open-source cloud distribution. It integrates components required for grid site virtualization and dynamic scaling to remote clouds, addressing the specific requirements of the grid resource providers, and for the deployment of science clouds, addressing infrastructure cloud-like access demands from user communities, including industrial users.

WP4 works in close collaboration with WP5 for production deployment, WP6 for new service and component integration and all other work packages.

WP4 is also responsible to the execution of the project agile process, which includes active participation from all work packages.

3.4.1 Summary

3.4.1.1 Quarter 5

QR5 was an opportunity, over the summer, to consolidate several features released in v1.0 and start development and integration work of several new exciting features. These range from persistence storage and better understanding of Amazon EC2 interfaces in preparation for hybrid deployments, but also better integration of Claudia client tools and general configuration.

The architecture for v2.0 was also reviewed and updated and documented in D4.4.

Work also took place on consolidating our continuous integration strategy with several new tests extending the test coverage to most services.

3.4.1.2 Quarter 6

During QR6, the integration and testing effort focused on the Persistent Disk Service, part of the StratusLab storage strategy. This important feature provides a significant boost in performance and scalability. The inclusion of OpenNebula 3.0 to the StratusLab distribution is also a focus, including several adaptations required on both sides to complete the integration.

The build and test infrastructure was improved by adding a certification staging step, with correspondent new infrastructure resources and jobs in our continuous integration server.

3.4.1.3 Quarter 7

During Q7, the integration and testing effort continued on the Persistent Disk Service, part of the StratusLab storage strategy. This service is sensitive to the local infrastructure network, which required further work in providing reliable behavior and acceptable performance. OpenNebula 3.0 was integrated in StratusLab, and integration of OpenNebula 3.2 started.

The improvements on the build and test infrastructure performed during QR6 payed-off, with significantly reduced certification time.

Three sprints were completed during this period, resulting in the release of StratusLab v1.2 and v1.3.

3.4.1.4 Quarter 8

During Q8, the integration and testing effort continued in order to deliver a solid v2.0 release of the entire StratusLab distribution. Several services were upgraded. We also completed the integration of OpenNebula 3.2.

The build and test system was also refactored to simplify it, in order to add support for release on CentOS 6.2, Fedora 16 and OpenSuSE.

Four sprints were completed during this period, resulting in the release of StratusLab v1.4 and v2.0 beta and v2.0.

3.4.2 Task 4.1: Definition of Reference Architecture

3.4.2.1 Quarter 5

The architecture of StratusLab v1.0 was updated to fulfill the needs of StratusLab v2.0. The main new focus of v2.0 is support for hybrid cloud deployment. This important task is ongoing and includes inputs from all work packages and partners.

The creation of D4.4 - Reference Architecture for StratusLab Toolkit 2.0, was also an opportunity to create service specific description, which will be used on our online documentation to provide better visibility and detail on our cloud distribution.

3.4.2.2 Quarter 6

The reference architecture of the StratusLab 2.0 distribution was discussed during the face to face meeting in Dublin, as well as during several TSCG conference calls, to ensure that it is still meeting the current project requirements and priorities. The addition of a jClouds client interface was added to the existing architecture.

3.4.2.3 Quarter 7

The architecture of StratusLab is stable and no significant modifications were required.

3.4.2.4 Quarter 8

The architecture of StratusLab is stable and no significant modifications were required.

3.4.3 Task 4.2: Integration of Open-source Distribution

3.4.3.1 Quarter 5

Persistent Storage A new important StratusLab feature is the support for persistent storage. While the foundations for such feature were developed and integrated during year 1, this quarter say this feature used and exposed to a number of end-user feature, notably the ability to create new persistent disk, which can be attached to virtual machines via standard commands.

Integrated Improvement from Operations This work also included integration of code developed by WP5 to improve the performance of the reference infrastructure, back into the mainstream StratusLab distribution, including testing and

installation support.

Development of StratusLab client tools for Claudia The development of StratusLab client tools for Claudia has started in this quarter. They are a set of python scripts, integrated with the StratusLab command line interface (CLI) tools, used to invoke the Claudia API, to deploy a service, undeploy it, obtain information about its status, and so on.

Code for RESTful Tests SixSq donated to the project a REST test framework, developed for SlipStream, such that StratusLab RESTful services can be better unit tested. Integration of this framework with services, such as the Marketplace has started, which should result in a more robust service with high test code coverage.

OpenNebula OpenNebula v3 is being actively developed, which includes StratusLab required features such as group support.

3.4.3.2 Quarter 6

In order to improve the pace at which we can release StratusLab, while reducing the effort in upgrading reference infrastructures and improve quality, a certification stage was introduced between the snapshot package generation and testing and the production infrastructure. To support this, WP5 deployed new machines dedicated to the certification tasks, such that stable packages can be deployed and tested, without being impacted by new snapshot packages being produced. To improve the deployment of StratusLab on this new testbed, new Hudson jobs where created in collaboration with WP5 such that these are automated.

A set of jobs is under development to generate, on-demand, new release package repositories, which contain release packages.

The integration of the Persistent Disk Server was significantly more complex than expected. Several important updates were required to the service implementation. Further, new OpenNebula instantiation scripts were required to integrate the Persistent Disk Server as a cache engine and quarantine. The image creation process was also significantly updated. The result is a significant boost in performance, with for example near instantaneous image instantiation and a five fold speed-up factor for image creation.

Development of StratusLab client tools for Claudia The StratusLab client tools for Claudia has been integrated in the StratusLab distribution in this quarter. They are a set of python scripts, integrated with the StratusLab command line interface (CLI) tools, used to invoke the Claudia API, to deploy a service, undeploy it and obtain status information.

Integration of monitoring systems components in StratusLab Monitoring components software have been integrated in StratusLab in the monitoring-code git repository. Two new RPM packages have been created for installing the collector service and the probes. Finally, new jobs in Hudson have been created for the inclusion of the software in continuous integration.

Development of admin tools for Monitoring systems The development of admin tools for monitoring systems, which are being developing in WP6, has started in this quarter. They are a set of python scripts, integrated with the StratusLab command line interface tools, used to install the different components involved in monitoring systems (database, API, collector, probes and so on).

Integration of OpenNebula 3.0 to the StratusLab distribution OpenNebula 3.0 is being included to the StratusLab distribution. Several adaptations are required on both sides to complete the integration, including changes in the command line interface, changes in OpenNebula to refer to objects by name, as well as configuration generation by StratusLab tools.

3.4.3.3 Quarter 7

Multiple OS Support Work to support more operating systems, in addition to the standard Fedora 14, has taken place. The target operating systems include OpenSuSE 12, Fedora 16 and CentOS 6.2. This ongoing work is significant, with important upgrades to our build and test system.

Integration of Monitoring API in StratusLab Monitoring component software was integrated in StratusLab in previous quarters. In this quarter, an implementation for a monitoring API (TCloud) has been introduced. This implementation is part of the tcloud-server-rpm.

Updated Claudia Client Tools The StratusLab client tools for Claudia have been modified to provide configuration for the monitoring systems.

Claudia Integration with OpenNebula 3.0 OpenNebula 3.0 was integrated in the StratusLab distribution, requiring several components to be adapted, including the clients and Claudia.

Port Address Translation PAT developed by IBCP was integrated; this allows virtual machines with local IP addresses to communicate with the outside world, permitting a site with a limited number of public IP addresses to still offer a viable public cloud service.

Persistent Disk Service The image creation was upgraded to take advantage of the pdisk service. The StratusLab TM uses store/snapshot capabilities of pdisk for caching of new images, fast start up of VMs, and easier, asynchronous new image creation and registration. This integration also includes a much more thorough policy check based on metadata in the Marketplace.

Several performance and stability issues were detected during this quarter. They now seem to have being addressed but this will have to be confirmed during the last quarter of the project.

3.4.3.4 Quarter 8

Placement Decision Module The placement decision module, developed in WP6, which is required for federation, has been included as part of the Claudia distribution in StratusLab.

The Persistent Storage Service was upgraded with a new backend architecture built in Python. This provides a cleaner separation of concerns between the Java Restlet service implementation, responsible for disk state, from the logic manipulating the backend storage system. The result of this separation of concerns also includes a cleaner integration with OpenNebula drivers. The Persistent Storage Service now supports iSCSI TGTD (as before) and iSCSI NetApp. Further, a File based backed is also available, and used at GRNET, making the Persistent Storage Service able to work with distributed file systems, as well as iSCSI solutions. We also have prototyped an iSCSI and Fiber Channel IBM V7000 implementation, although this prototype is not production ready.

OpenNebula 3.2 is now integrated with StratusLab. Several backward incompatible changes were detected during the integration, but only after the release of 3.2, which forced us to adapt and spend more effort than expected.

Version 2.0 of StratusLab is now being released under CentOS 6.2 as the default platform. This will help enterprise adoption of the StratusLab distribution, since compatible with Redhat Enterprise Linux 6.2 (RHEL 6.2).

3.4.4 Task 4.3: Contextualization of Grid Services

3.4.4.1 Quarter 5

Contextualization was improved via richer metadata creation using the StratusLab command line tools.

A deprecation feature was also defined, in collaboration with WP5 and WP3, including new commands as part of the existing command line client.

3.4.4.2 Quarter 6

WP3 provided support to WP5 in the creation of automatic deployment of Torque-based clusters using SlipStream on the StratusLab reference infrastructure at GRNET. This work provides the EGI/EMI community with an easier access to cloud resources for common computing patterns.

3.4.4.3 Quarter 7

With the more stringent policy enforcement feature integrated in StratusLab v1.3, several Marketplace entries had to be corrected and improved in order to be compliant.

WP3 provided support to WP5 in the creation of automatic deployment of Grid Sites using SlipStream on the StratusLab reference infrastructure at GRNET. This work, although not completed yet, will provide the EGI/EMI community with a simpler way of deploying and configuring a Grid Site on the StratusLab cloud.

3.4.4.4 Quarter 8

An initial discussion took place around Cloud-init, proposed by Ubuntu, and endorsed by Amazon, which proposes a first convention for contextualisation. We have not had the chance to converge on this interesting development, but this discussion will continue, with a potential implementation, after the project's conclu-

sion.

The promising element of cloud-init is that it can be implemented on top of the current StratusLab contextualisation strategy, confirming that our initial assumptions and strategy were right.

3.4.5 Task 4.4: Technical Support

3.4.5.1 Quarter 5

Claudia Technical Support An important effort has been done during this quarter for providing technical support for Claudia. In a way, several bugs in the Claudia code have been solved and some improvement has been carried out in the generation of packages (poms, hudson jobs etc.) for version 1.0. In addition, technical support is being provided to StratusLab end-users for installation, configuration and usage of Claudia.

Technical support was provided, tightly coordinated with WP5, on a wide range of issues, such as troubleshooting virtual image contextualization and command line tools semantic.

3.4.5.2 Quarter 6

Technical support continued to be provided, coordinated with WP5, on a wide range of issues, including the Persistent Disk Service, contextualisation and command-line tools.

3.4.5.3 Quarter 7

Technical support continued to be provided, coordinated with WP5, on a wide range of issues, including the Persistent Disk Service, contextualisation and command-line tools.

3.4.5.4 Quarter 8

Support for Integrating OpenNebula 3.2 A new set of configuration files has been created for StratusLab installer. Code has been adapted to make OpenNebula 3.2 behave more similarly to 3.0. Patches previously done for 3.0 have been translated to work with 3.2. Finally, the CLI has been also changed to comply with the new XML-RPC API.

New Hooks for User Management New hook functionality has been added to execute scripts on user creation or modification. This was needed for the stratuslab-one-proxy and the SunStone web interface integration.

Integration of the SunStone Web Interface The SunStone web interface for OpenNebula has been integrated in StratusLab, to graphically show the deployment of VMs in demonstrations.

Preparation of a Scenario for Cloud Bursting An scenario has been prepared where OpenNebula provides access to Amazon EC2, using its adapter. In OpenNebula, a public Cloud service like this is managed as any other host, but it may provide “infinite” capacity for the execution of VMs. Uploading VM images on-

the-fly to the external cloud is still not feasible (because the Amazon S3 datastore is not ready yet), therefore it is assumed that a valid VM image has been already uploaded to EC2. And, since VMs deployed on different clouds have to communicate through the Internet, OpenVPN is used to establish a secure channel between them.

Preparation of a Scenario for Cloud Federation Another scenario has been prepared where OpenNebula provides access to other StratusLab sites running a different OpenNebula instance. Other StratusLab sites are accessed using the ONE2ONE adapter. For storage, the sites just use the same Markeplace service. For networking, the solution based on OpenVPN is applied, as in the Cloud bursting scenario.

Support was also provided for a number of deployments, including at GRNET and LAL, but also of evaluation infrastructures in Vietnam and for the Helix Nebula collaboration.

3.4.6 Issues and Corrective Actions

3.4.6.1 Quarter 5

A higher number of build and test job failures were recorded during the second part of the quarter. This is in part due to the holiday period were fewer people were available to quickly fix problems as they occur. This is also due to the high rate of code commits performed around the new features for persistence storage and Claudia tools. This had no impact on production. Specific tasks were created in Sprint 17 to address this issue, to return our continuous integration strategy to a nominal state.

3.4.6.2 Quarter 6

LAL experienced a severe loss of cooling in its datacenter for a period of two weeks. This impacted seriously the StratusLab build and test infrastructure. This meant that continuous build and test jobs could not run during that time. During this time, technical tasks continued with regular updates to code and configuration, with however without feedback from the build and test jobs normally immediately alerting team members of problems after commit. This is in large part responsible for the fact that only one release of StratusLab was performed during this quarter. The build and test jobs are now running without problems, but this came at a significant effort, to stabilize the build and test jobs. In the process, we improved the job structure, simplifying it and improving its performance (e.g. faster build and test) and introduced the certification and release jobs to reduce manual tasks in certifying new StratusLab release candidates. This should therefore improve the release rate for the rest of the project.

3.4.6.3 Quarter 7

The behavior of pdisk continued to cause problems to the StratusLab reference architecture. Several improvements were investigated and implemented, which have

significantly improved the situation. We believe that the latest solution provide reasonable performance and stability. This effort in testing and monitoring of the storage solution will continue during the last quarter of the project.

3.4.6.4 Quarter 8

No serious issue encountered during this quarter.

Backward incompatible changes were detected with OpenNebula v3.4. The changes relate to the fact that image registration with the OpenNebula image repository is now made mandatory. This means that StratusLab is not currently compatible with the latest OpenNebula version. A possible way forward would be to build a rich facade to the OpenNebula Daemon, hiding some of these incompatible changes. However, this was not possible within the effort and time constraints left in the project. This is an issue that will have to be carefully handled as we transition towards the community driven open source project.

3.5 WP5: Infrastructure Operation

WP5 is responsible for the provision and operation of the project's computing infrastructure. It serves as a beta-tester of the software integrated by WP4 and WP6, deploying it in a production environment in order to verify its applicability for real life applications. In addition WP5 offers daily support to external users, either system administrators or those exploiting the project's public cloud services. Finally, WP5 is contributing with targeted development activities, related to the improvement of the cloud tools and services, as well as to showcase the capabilities of the StratusLab distribution to satisfy different use cases.

3.5.1 Summary

3.5.1.1 Quarter 5

WP5 started the Y2 activities having two clear goals; continue the successful operation of public cloud services by constantly integrating and improving the StatusLab distribution and improve on the areas identified during the first year project review as being weak. The recommendations of the review have been taken promptly into consideration and in the past quarter and have been the driving force for two main activities:

- Expand the scope of target use cases demonstrating the capabilities of the software to satisfy a multitude of applications.
- Improve the performance of the offered public services, mainly in the area of storage management.

3.5.1.2 Quarter 6

During Q6 work on Operations focused on the redesign the pre-production cloud services and the expansion of the resources providing the reference cloud service. In order to better streamline the adoption of the software developed by WP4 (an issue that has already been identified in the past reports) a new software certification process has been put in place and is being finalized. The main tools of this process will be the new certification pre-production site, the certification yum repository and a set of Hudson jobs that automate the deployment and test of release candidate versions of StratusLab software packages. Finally on the problems side, there have been some issues with the availability and reliability of the production grid site and the physical infrastructure hosted at LAL.

3.5.1.3 Quarter 7

Evolution, testing and fine tuning have been the driving forces for WP5 during this quarter. The activity followed closely the evolutionary path of the StratusLab distribution by certifying and installing all the releases of the software during this quarter. During this process we intensified the effort required for testing the new releases. We also identified issues in the software and in the deployed hardware

architecture, trying to alleviate them (at least those issues related to the physical infrastructure). For the hosted virtualized grid site, there is an ongoing effort to migrate to the new software supported by EGI (UMD-1) and to provide prepared appliances and support tools in order to ease the deployment and configuration of grid services. Improvements will be made in all in the coming months, correcting some unresolved issues pertaining the above tasks.

3.5.1.4 Quarter 8

The focus of work for WP5 during Q8 evolved naturally around the fine tuning of services and the resolution of outstanding issues in order to deliver the final version of services and the overall activity results. Among the highlights of this last quarter is the introduction of common authentication service between LAL and GRNET sites and the testing and deployment of file-based persistent disk service.

3.5.2 Task 5.1: Deployment and Operation of Virtualized Grid Sites

3.5.2.1 Quarter 5

Production Cloud Service The reference cloud service was upgraded to StratusLab v1.0 immediately after the official release of the software. The interference of the upgrade was limited since the service was already running on v0.4 of the distribution which incorporated most of the new features introduced to the software in the last months (including the switch of the hosting OS from CentOS 5.5 to Fedora 14).

Operation of a Production Grid Service The grid site was reconfigured in order to support more EGI VOs. Thus by the end of Q5 HG-07-StratusLab supports the following scientific VOs:*alice, atlas, biomed, cms, compchem, env.see-grid-sci.eu, esr, gridcc, magic, meteo.see-grid-sci.eu, nwchem.vo.hellasgrid.gr, planck, scier, seismo.see-grid-sci.eu, vo.complex-systems.ey* and *vo.dorii.eu*. The selection of the VOs to support was imposed by the policies and procedures defined by NGI-GRNET which is the hosting NGI were the grid site is located. By supporting these VOs, StratusLab's production grid site provides computing resources to scientific domains like HPC, Bioinformatics, Biomedicine, Computational Chemistry, Seismology, Meteorology, Complex Systems, Computer Science and Astrophysics.

In parallel, the total number of computing resources were doubled in size in order to better accommodate the expected increasing demand from the new VOs. The current total amount of CPU cores offered by the site are 32 hosted in 16 virtualized WNs running on the StratusLab reference cloud service.

Adaption of MapReduce platform In order to demonstrate the applicability of StratusLab to satisfy different use cases, we adapted the well-known Hadoop platform and demonstrated the ability to run MapReduce applications. MapReduce (http://en.wikipedia.org/wiki/MapReduce) is a patented software framework introduced by Google in 2004 to support distributed computing on large data sets on

clusters of computers. Apache Hadoop (http://hadoop.apache.org/) is an open source implementation of MapReduce. In order to demonstrate this capability we produced a Virtual Appliance based on CentOS 5.5 that comes pre-installed and configured with *Hadoop 0.20.203* and *JDK 1.6*. This appliance is available from StratusLab's Marketplace (image ID: *Gz3B6D3z1F1m4k6RXaQA_gc0nop*). The instantiation of the Hadoop cluster takes advantage of the *stratus-run-cluster* command as documented in the project site: http://stratuslab.eu/doku.php/tutorial:mapreduce.

Following this effort we continued the tests with Hadoop by exploiting the SlipStream platform as a tool to instantiate, configure and run a sample MapReduce application. For this purpose one additional virtual appliance was produced, based on Fedora 14. This was enforced due to the requirements imposed by SlipStream on the virtual images, that have to support Python 2.6 or later. The appliance has been configured similarly to the CentOS one and is available from the marketplace with the ID: *FcQdN7cOjfueaFmb-aJ80_TqHxJ*.

Currently, we plan to demonstrate both of the above activities during the EGI Technical Forum in Lyon (Sept 2011).

Improved Contextualization The contextualization framework was improved making it more flexible (especially for automated installations) and packaging it to make it easier to install in RedHat-based distributions.

3.5.2.2 Quarter 6

Expanded physical resources for reference cloud service The physical infrastructure providing the StratusLab reference cloud service that is located in GRNET has been expanded with the addition of two more physical nodes. Thus the total resources provided now are 18 hosting nodes plus the front-end.

Reference cloud service upgrade to StratusLab 1.1 Shortly after the public release of v1.1 the reference service was upgraded to the latest version. In principle the set of services provided has not changed with the only exception of Claudia Service Manager which is now properly configured and available from the service front-end.

Expanded IP address range The range of public IP addresses that are offered to VM instances from the DCHP server has been expanded by 135. This means that now all the 62.217.122.0/32 subnet has been dedicated for VMs. As a result of this upgrade, lack of public IPs is no longer a limitation for the reference cloud service utilization.

Migration to new DNS domain The infrastructure has moved from the rather obscure **.one.ypepth.grnet.gr* domain to a more clean **.grnet.stratuslab.eu* domain. The primary domain is managed by LAL and the "grnet" subdomain is managed by GRNET itself. This led to a change of the official hostname used for accessing the cloud front-end which now is cloud.grnet.stratuslab.eu (was cloud-grnet.stratuslab.eu). All the registered users had been informed of the change.

Redesigned support infrastructure The pre-production service has now been replaced by two individual testbeds: a development testbed and a certification testbed. The development testbed comprises two physical nodes and will be used for testing, experimentation of new features and overall development of new services. The development testbed typically runs a development version of the StratusLab distribution. The certification testbed also comprises two physical nodes (one fronted and one hosting node). The latter testbed will be used to test and certify the release candidate versions of StratusLab distribution. An automated procedure based on Hudson is currently under development (see testing section).

Development of the Torque cluster use case In the context of our collaboration with EMI there was a request to provide Torque ready VM appliances; essentially a stripped-down version of grid cluster without the rest of the grid software. As a result a new appliance has been developed and tested and is available from the Marketplace[21]. A single image is used to instantiate both the server and worker part of the cluster. A script that is available inside the image can be used to automate the installation process. Instructions on how to instantiate the cluster using the stratus-run-cluster command is available from the project web site[22]. In addition a StratusLab deployment module was developed as an alternative, more generic solution for cluster instantiation.

Refactored Persistent Disk Storage The persistent disk storage service was significantly refactored to make it more efficient, reliable, and testable. This was done to make it easier to implement image caching features developed by WP4.

Service Certificates for stratuslab.eu Domain The stratuslab.eu domain was transferred officially to CNRS allowing Terena certificates to be obtained for the stratuslab.eu domain. This benefits users of the StratusLab services because these certificates are trusted by default in all major browsers, avoiding spurious security warnings when using our services. These certificates will be rolled out progressively to all of the StratusLab public and private services.

3.5.2.3 Quarter 7

Reference Cloud Service The reference cloud service has followed the evolution of the StratusLab distribution, deploying first in a large scale environment the software delivered by WP4. In particular during Q7 the service has been upgraded to versions 1.2 and then 1.3 introducing in parallel the persistent disk service (pdisk) thus providing storage management and compute services. Also, the reference service was able to exploit the caching mechanisms (also a feature of pdisk) reducing significantly the amount of time it takes for a VM to instantiate improving the overall efficiency. Cached images take only a few seconds for the VM to reach the 'running" state, with the latency for access by the user dominated by the configuration of the image (number of default services and kernel modules).

[21] Identifier: AViUGHkhd5NytPAKnIp03hnsVWG

[22] http://stratuslab.eu/doku.php/tutorial:torquecluster

The transition to pdisk was not seemless though. The introduction of the dependency on iSCSI had a significant impact on the way storage is managed from the service backend. To date, the service utilized a centralized storage server (EMC Celerra NS-480) in order to provide a shared repository between the cloud frontend and the hosting nodes for the VM images. This setup was changed when deploying StratusLab 1.2. In particular an NFS volume was allocated to the frontend in order to host a large LVM volume which was then assigned to the pdisk server in order to offer disk devices over iSCSI for the VMs. This setup proved to be inefficient, introducing multiple levels of delay, degrading the overall performance of the service and in the end wiping out the expected benefits of pdisk. To work around this problem, we decided to drop the usage of NFS and alternatively use a relatively large local disk in order to configure the necessary volume groups for pdisk. This reconfiguration improved slightly the performance of the system with respect to storage I/O but still is underperforming comparing to the initial NFS-based setup. For example the I/O throughput recorded in the local disk iSCSI setup was around 5 MB/s whereas the throughput with the NFS is around 10 MB/s. Though both numbers are sub-optimal the difference between the two approaches is clearly evident.

Finally another problem identified was the degraded performance of the existing appliances that have been prepared using the qcow2 format. Qcow2, as the name suggests, uses a copy-on-write (cow) approach, starting from a minimum disk space and expanding to a maximum predefined size at run time based on the image disk usage. Unfortunately this behavior introduces one more performance bottleneck, since the images are used to create an LVM volume that performs poorly during copy-on-write activities.

Hosted grid site The production Grid site running on the reference cloud service, is scheduled to follow the evolution of EGI software and migrate to UMD-1 software distribution. For this purpose a set of appliances has been prepared with the required software for the CE, SE, WN and APEL grid nodes. In parallel a deployment module has been developed using Slipstream in order to automate the grid site configuration procedure. By the end of Q7 the transition has not been completed yet due to the performance issues described above. The site is still running on a backup cloud service that does not use pdisk till the issues of the reference cloud have been solved. The instantiation of the site with the UMD-1 images are expected to complete at the beginning of Q8 during Sprint 23.

Persistent Disk Service Improvements to the persistent disk service were made to make it more stable and more modular. This modularity allows different storage options (posix file system, LVM) and sharing options (iSCSI, shared file system) be used by the service. A prototype module has been written that allows a NetApp storage system to be used behind the persistent disk service. This will be throughly tested and put into production in Q8.

TERENA Certificates Work was done to ensure that all of the StratusLab services have valid certificates issued through TERENA. These certificates are signed

by a standard Certificate Authority and do not require special configuration of clients or cause warnings about untrusted certificates to appear in browsers.

3.5.2.4 Quarter 8

GRNET Reference Cloud The site continued to operate during Q8 using the iSCSI-based pDisk service running on the frontend node. After the upgrade to StratusLab 1.4 the behavior of the service is much more stable and the only real issue is the lack of storage space in order to deploy a large number of VMs. With 1 TB local storage the maximum capacity of the site is around 30-40 VMs depending on the volume of cached images. This restrictions have motivated the development of the file-based pDisk version which is planned to be included in StratusLab 2.0. As of the writing of this report, we are currently testing this version and we plan to deploy it in the production site once certified.

LAL Reference Cloud The cloud infrastructure at LAL has been opened up to external users. This serves as an alternate site to the GRNET infrastructure and allows for testing of different storage solutions. (LAL has both storage that can be accessed via an iSCSI target like tgtd and a commercial NetApp solution.)

Unified Authentication The two reference cloud services now use a single LDAP server for authentication. This allows users to access both sites using the same credentials. This LDAP server is also connected to a simple registration service that allows users to register for the StratusLab cloud and manage their own password and certificate information.

Hosted grid site HG-07-StratusLab continued to run on an an auxiliary cloud site in order to mitigate the restrictions of iSCSI-based pDisk. The site will also be moved to the production cloud service in GRNET once the upgrade to v2.0 is completed and the site starts exploiting the new file-based pDisk service.

3.5.3 Task 5.2: Testing of the StratusLab Toolkit

3.5.3.1 Quarter 5

Improving the cloud service features and performance Two activities were carried in the pre-production service; the development of an NFS-based volume management service similar to EBS (Elastic Block Storage) and the introduction of a caching mechanism in order to speed up the VM instantiation process.

The NFS-based volume management service exploits the storage space provided by NFS in order to create virtual machine disk images and attach them to running instances. The functionality has been installed on the pre-production cloud service in GRNET and was demonstrated during Sprint 16 and is currently in the process of being integrated to the release code.

The image caching mechanism has been introduced in order to tackle the latency issue that we experience during VM instantiation. This latency is introduced by the fact the each time a VM is instantiated the respective image has to be transferred from a remote repository (typical from the site in TCD). The caching mecha-

nism utilizes the capabilities offered by QEMU and uses a combined configuration of SSH and NFS installation on the cloud site that permits copies of VM images to be stored locally, shared among all hosting nodes and used whenever a new VM is instantiated. The mechanism has been developed and demonstrated also on the pre-production site in GRNET and we plan to integrate it in the stable version of StratusLab distribution at the beginning of Q6.

3.5.3.2 Quarter 6

Automated certification of release candidate versions The goal of this task is to streamline the interaction between WP4 and WP5 and in particular the development, certification, and deployment process for new versions of the StratusLab distribution. In order to enable this, three components have been created:

- A certification yum repository that contains release candidate versions of rpm packages comprising the next version of the software distribution,
- A certification testbed, described above. Provides the necessary physical resources for performing the deployment and verification tests, and
- A set of Hudson jobs that execute the necessary installation steps automating essentially the work that was done manually up to now in the former pre-production testbed.

The procedure is currently being finalized and will be put into action for the certification of version 1.2 expected near the beginning of Q7.

Storage Resources for LAL Cloud Service Over this quarter various improvements to LAL's cloud services have been done. In anticipation of opening the resources as a production service, a large amount of storage (40 TB) was acquired, half as a standard SAN and the other as a iSCSI server. The SAN has been integrated into the platform; the iSCSI server will be integrated in the next quarter. This will allow performance comparisons between the two systems.

Improved Monitoring for LAL's Resources LAL has upgraded all of its machines to use IPMI. This will allow more detailed monitoring of the resource utilization. In addition, a subset of machines has plug-level monitoring of the electricity consumption, allowing a cross-check of the IPMI values. This should be fully functional in Q7 and be able to provide feedback on resource utilization with different cloud configurations and services.

3.5.3.3 Quarter 7

Automated Certification of Release Candidates The new certification process defined in Q6 were utilized during Q7 in order to test and verify versions 1.2 and 1.3 before releasing them and deploying them in the reference cloud infrastructure. As expected the automated process significantly accelerated the software validation procedure. Nevertheless it has not yet reached to a level of completeness that will enable the identification of all potential issues in a release. For example the process failed to identify the performance bottlenecks that appeared in

the reference deployment. The problem arises from the fact the the certification is performed on a limited infrastructure (1 frontend - 1 hosting node) which is not sufficient to perform large scale tests (e.g. bulk instantiation of 30 or more VMs in order to stress the system) nor can offer all possible hardware configuration scenarios (e.g. using centralized storage server or a distributed file system). During Q8 we will try to improve this situation by providing additional resources (even temporarily during the certification) in order at least to be able to identify any potential scaling issues.

Support for virtio_net and virtio_blk Drivers This activity was initiated from a question from one of our users noticing that the maximum bandwidth that he could achieve on the reference cloud was 100Mb/s. Initial experiments revealed that actually the maximum bandwidth that can be achieved is around 150 Mb/s. Having in mind the the hosts are using a 1 Gb/s Ethernet interface this number is significantly lower than the maximum theoretical peak. Further investigation revealed that the cause of the problem is that the VM appliances are not configured with the virtio paravirtualized driver. Therefore, the network requests from the guests are emulated, which results in poor I/O performance. By activating virtio_net on both VM and hosting node the network bandwidth increases up to 680 Mbit/s.

Similarly, the access to the local disk passes through the native IDE driver and not through the paravirtualized virtio_blk one, although the performance difference in this case does not appear to be significant.

The above issues have been forwarded to WP4 as a requirement and is expected that the next versions of StratusLab will offer native support for virtio_net and virtio_blk.

3.5.3.4 Quarter 8

NetApp Support Integration of existing resources at a site is important for the adoption of the StratusLab cloud distribution. This is especially important for storage as sites are very likely to have their own storage infrastructure already in place. LAL has worked to integrate its NetApp server into the persistent disk service. This has resulted in a refactoring of the persistent disk "backends" to make the integration of new services more flexible. This work was also necessary to allow shared file systems to be used with the persistent disk service again.

IPv6 Support With the exhaustion of the IPv4 address space, support for IPv6 is becoming more important. The StratusLab services have been tested in a mixed IPv4 and IPv6 environment. All of the java-based services (the majority) worked without any changes in this environment; those services were accessible to both IPv4 and IPv6 clients. OpenNebula cannot at the moment provide both IPv4 and IPv6 addresses to clients. This is not really a limitation in the StratusLab configuration as IPv4 can be used in OpenNebula to assign addresses and the site DHCP server can be configured to provide both IPv4 and IPv6 addresses to the associated MAC address. A slight change was required in the StratusLab client to provide the users with domain names rather than raw IPv4 addresses. It was also required to

update the standard base images to configure both IPv4 and IPv6 networking.

File-based pDisk service As mentioned above, the restrictions imposed by the LVM/iSCSI based pDisk service has led to the development of a file-based approach which can exploit any local or network file service for hosting and sharing VM instance images and storage volumes. The "file-based" term refers to the fact that VM images are simple files and not logical volume devices as in the iSCSI implementation. This approach not only enables the utilization of a large range of file systems (NFS, GPFS, Ceph, etc) but maintains the original file format for VM instances which in turn is very useful for features like instance live-migration, VM backup etc. This version of pDisk is currently under certification and will be released officially as part of StratusLab v2.0 by the end of the project.

3.5.4 Task 5.3: Virtual Appliances Creation and Maintenance

3.5.4.1 Quarter 5

Marketplace In this quarter the focus of the task has been on the continued design and implementation of the Marketplace. The Marketplace has been made available for use at http://marketplace.stratuslab.eu. The Marketplace has now replaced the appliance repository as the central location for sharing of image metadata. The appliance repository is maintained to provide storage resources to the project for the actual image files. The Marketplace interface has been improved to provide easier browsing and searching of metadata entries. Work has also started on implementing a deprecation mechanism for metadata entries that is compatible with the EGI draft policy on the endorsement and operations of virtual machine images.

Existing Hudson jobs have been extended to further test the deployment, configuration, and functionality of the Marketplace. These run regularly and flag any errors that have been introduced in the code or the deployment procedures.

3.5.4.2 Quarter 6

Marketplace During this quarter development of the Marketplace has continued. The focus has been on improving testing, scalability and usability. Unit tests and resource-level testing have been provided. Processing of the metadata display table has been moved to the server-side. The implementation of the deprecation mechanism has been completed. A new client command *stratus-deprecate-metadata* has been provided, which allows users to easily deprecate Marketplace entries.

3.5.4.3 Quarter 7

Marketplace During this quarter development of the Marketplace has continued. The focus has been on improving documentation and performance. Documentation on installation and usage has been added to the project wiki.

Support for PostgreSQL as a back-end for metadata storage has been added. This provides significantly better performance than MySQL.

A requirement of the Marketplace is that email verification should be per-

formed when metadata is uploaded. To date this has not been enabled as it would prevent automated uploads. To allow for services to upload metadata the ability to whitelist certain endorsers has been added.

Automated Production of Base Images Work was done, in collaboration with WP2, to provide a consistent procedure for creating the StratusLab base images. Moreover, this production was automated to ensure that security patches can be released regularly. See the more detailed description in the WP2 section.

3.5.4.4 Quarter 8

Marketplace During this quarter development of the Marketplace has continued. The focus has been on bug fixing and interface improvements.

Marketplace availability is critical to the StratusLab infrastructure, and as such a replication mechanism is required to provide fault-tolerance in case of failure. A simple scheme has been implemented that will allow for a replica Marketplace to be quickly and easily deployed using standard tools (rsync, and cron). A replica Marketplace was deployed at TCD to demonstrate this.

A number of fixes were made for the email verification, which is now in use at IBCP for their Bio-Informatics Marketplace. Also, some improvements have been made to the search interface to make it easier for users to filter entries. Search performance has been improved through changes to the internal SPARQL queries.

3.5.5 Issues and Corrective Actions

3.5.5.1 Quarter 5

Availability and Reliability of production Grid site The average reliability and availability metrics for HG-07-StratusLab in Q5 exhibit slightly lower numbers than the foreseen target for Y2. In particular the average availability was 92% and the reliability was 91% (both targets for both values are ¿95%). The main reason for these degraded values was the major upgrade that we performed during June that required the complete re-installation of the reference cloud service physical hosts. The operation system of the hosts had to be switched from CentOS 5.5 to Fedora 14. Obviously such a radical upgraded required a few days of down-time in order to backup the images, re-format the nodes, install StratusLab and re-instantiate the VMs in the new setup. This is reflected also in the monthly statistics of June in which reliability and availability were 82% and 79% respectively. During July and August the numbers returned back to the regular high level thus increasing the overall average.

The above figures first of all underline an obvious fact that the performance and quality of a virtualized service depends heavily on the underlying cloud platform. Moreover it's a clear indication that one of the things we should improve in the operations part is the process of core software update. This process should be as smooth as possible limiting the time that the service is brought of line. For this goal we will need the combined effort from WP4 activity which has to ensure that the core components of the cloud service (VM management, storage, network etc.)

are able to migrate to new versions of the software fast and with limited manual intervention from the cloud administrators.

3.5.5.2 Quarter 6

Degraded availability and reliability numbers for the production Grid site The availability and reliability numbers have been considerably affected in the past quarter due to problems with the site BDII service on the CE. The latter has been hanging regularly with no apparent reason making the site unusable for the rest of EGI infrastructure. Most probably this is a problem of the grid software and not of the underlying cloud service but the issue is still under investigation. For the time being we have tried to overcome the problem with closer monitoring of the service and the quick intervention from administrators to bring the service back online.

Slow adoption of new features by the reference cloud service Many of the technologies developed by WP4 still have not been adopted by WP5 including persistent storage and image caching. Will try to catch up in release 1.2. New features are difficult to adopt immediately since in many cases they bring radical changes to the cloud service and require service redesign and long downtimes. Overall, this is the typical development-operations problem that we try to solve with the automated certification procedure.

Delay in delivery of D5.4 The deliverable D5.4 titled *"Economic analysis of infrastructure operations"* which was planned for PM18 has to be delayed until PM21. The delay has been requested in order to collect more data from the operation of the production grid site which has been running only a few months in full operation (since July 2011). We believe that this extension of the delivery date will serve better the purposes of the deliverable since it will give us more time to perform a thorough analysis of the data and thus provide better conclusions regarding the actual economic impact of grid site operation over cloud services.

Cooling Outage at LAL There was a significant outage of cooling in the LAL machine room for a few weeks around the beginning of November. This negatively impacted the standard build and test services and delayed anticipated releases of the StratusLab toolkit. It also greatly reduced the resources available for a StratusLab tutorial in Vietnam. This has since been corrected and hopefully there will not be a recurrence of this problem.

3.5.5.3 Quarter 7

Performance of the reference cloud service As mentioned already with the introduction of pdisk we experienced a performance degradation for what concerns activities that depend on the disk I/O. From our tests it is evident that the way pdisk works (LVM with iSCSI) requires a different hardware setup than what we currently have in the reference infrastructure. In particular the service performs on par when a dedicated iSCSI server or storage appliance is used to serve volumes. This setup for example is currently employed successfully in LAL's installation. In order to alleviate this issue in GRNET (where the reference infrastructure is

located) we are currently investigating alternative setups that will resolve the bottlenecks experienced in the current configuration. Meanwhile, since this situation has impacted our work on the production grid site we are thinking to move the latter from the GRNET cloud to LAL's cloud. This will also give as a good opportunity to test and demonstrate the federation capabilities of StratusLab and also the ability to move virtualized grid sites between different cloud services, even different countries, still maintaining the quality levels required by EGI.

3.5.5.4 Quarter 8

Security incident in SlipStream nodes One of the physical nodes in GRNET that hosts a production instance of SlipStream was hacked and used to launch further attacks to VNC servers running in other hosts. The incident was brought to the attention of the WP5 operations team by GRNET's Network Operations Center. A forensic analysis in the node revealed a rootkit installed under /tmp/.ICE-unix/.run containing all the necessary scripts to take control of the local system and initiate the attacks to VNC ports of external systems. The security break was tracked back to host with IP 85.183.72.18 located in Germany. After the forensics completed, the node was re-imaged and returned to SixSq in order to re-install SlipStream. GRNET's NOC was informed regarding the above findings for further actions.

Grid site elasticity The ability to deploy elastic grid sites that autonomously change their size using Claudia's KPI capabilities has been demonstrated in previous quarters. Nevertheless, this features was never moved to production usage. This is because we finally decided to give priority to other more important issues like the testing and feedback of the file-based pdisk services as well as the optimization of the existing static grid site and the migration to UMD-1.

3.6 WP6: Innovative Cloud-like Management of Grid Services and Resources

The Joint Research Activity (JRA), carried out in WP6, develops advanced technology and features for deployment on existing Cloud infrastructures through automatic deployment and dynamic provision of grid services as well as scalable cloud-like management of grid site resources. More specifically, the objectives to be accomplished can be expressed as: i) the extension of currently available open-source service-level frameworks which provide elasticity on top of cloud infrastructures, ii) the invention of new techniques for the efficient management of virtualized resources for grid services and iii) the inclusion of novel resource provisioning models based on cloud-like interfaces.

3.6.1 Summary

3.6.1.1 Quarter 5

The work done in QR5 mainly has involved the resolutions of bugs for the first review demo and the continuation of some missing tasks for that. In addition, some new functionalities have been implemented as the inclusion of new policies for scaling down and the management of the physical infrastructure just in the virtual machine layer. A new networking model have been included also in OpenNebula. Finally, some work has been done in authentication for the service manager API, the support for groups and roles, the development of a new VM Template Repository and the inclusion of new drivers for accessing to more Cloud providers.

3.6.1.2 Quarter 6

The work done in QR6 mainly has involved the definition of the work plan for WP6 as documented in the deliverable D6.4. Monitoring system components have been included and extended in StratusLab in order to obtain metrics from different layers. Some work related to networking has been done in this period, mainly network isolation through VLAN and the firewall management. Monitoring, accounting and authentication have been improved in OpenNebula and some support has been integrated for user data injection.

the first review demo and the continuation of some missing tasks for that. In addition, some new functionalities have been implemented as the inclusion of new policies for scaling down and the management of the physical infrastructure just in the virtual machine layer. A new networking model have been included also in OpenNebula. Finally, some work has been done in authentication for the service manager API, the support for groups and roles, the development of a new VM Template Repository and the inclusion of new drivers for accessing to more Cloud providers.

3.6.1.3 Quarter 7

The work done in Q7 has involved the continuation with the integration of the monitoring systems in the StratusLab distribution. This has implied to integrate monitoring systems with Claudia to drive with scalability and to introduce monitoring probes by contextualization. Regarding network management, the network operations are now coupled with the VM lifecycle, its definition is flexible and it is possible to reserve network leases. Placement policy can be defined in VM-basis too. Finally, the work on interoperability and APIs have carried out the introduction of the TCloud API as monitoring API, new OCCI extensions to include VM types and network definition, the introduction of the Deltacloud API and more attributes for Security and User Management. This development work is being integrated into StratusLab distribution in WP4.

3.6.1.4 Quarter 8

The work done in Q8 has involved working on Inter-Cloud Connectors for different scenarios. On one side a decision placement module has been created and integrated in Claudia for deciding in which site to deploy the service, providing Cloud brokering. On the other hand, Amazon EC2 and ONE2ONE drivers have been improved for having Cloud bursting and federation, respectively. In addition, Claudia has been evolved to include the persistent disk interaction. Finally, OpenNebula has been improved with support for multiple Datastores, new transfer drivers, or Cloud partitioning.

3.6.2 T6.1: Dynamic Provision of Grid Services

3.6.2.1 Quarter 5

Scale down strategy policies Due to grid service requirements, it has been created more scale down strategy policies. Now the node to be removed should be decided by the Compute Element, which knows the status of each worker node. It has implied the implementation of this policy in the service manager and the communication of service manager - CE in order to obtain this information.

Skip the IP management from Claudia In previous versions, Claudia had the management of networks and assigned IPs to virtual machines. In a way this is not a normal working mode, since the infrastructure should be managed by the virtual machine manager. Thus, Claudia has skipped this infrastructure managements and deploy the VMs in the networks created by StratusLab (public, private, local) where the DHCP assigns the right IP.

Service Manager - CE integration During the tests with site elasticity (i.e. the dynamic addition and removal of WNs) we've realized that the Local Resource Manager (Torque) could not cope with the continuous changes of the underlying infrastructure. The problem resided in the way torque communicates with the worker nodes and in particular with the setup of the ssh (Secure Shell) service. A new unix demon had to be introduced (wnMonitor) that continuously monitors the yaim con-

figuration file (wn-list.conf file) and whenever a new Worker Node is added in the cluster it updates the ssh host keys in the CE. This way it is ensured that torque server will be able to communicate without problem with the workers over the ssh protocol.

Service Manager - CE integration During the tests with site elasticity (i.e. the dynamic addition and removal of WNs) we've realized that the Local Resource Manager (Torque) could not cope with the continuous changes of the underlying infrastructure. The problem resided in the way torque communicates with the worker nodes and in particular with the setup of the ssh (Secure Shell) service. A new unix demon had to be introduced (wnMonitor) that continuously monitors the yaim configuration file (wn-list.conf file) and whenever a new Worker Node is added in the cluster it updates the ssh host keys in the CE. This way it is ensured that torque server will be able to communicate without problem with the workers over the ssh protocol.

3.6.2.2 Quarter 6

Design of the Cloud-like Management of Grid Sites The deliverable D6.4 has updated the component design defined in D6.1 with the work to be done in Y2. The main contributions of this report involve the inclusion of networking and image management functionality as well as the inter-cloud broker. In addition, Cloud APIs have been described in the interoperability and some updates have been included for the monitoring and accounting functionalities.

Monitoring systems The monitoring component in Year 1 has been updated to take into account multi-tier application requirements. That is, monitoring takes metrics from both VMs and hosts and incorporates probes inside the VMs to measure different parameters. Besides hardware information (RAM, CPU, etc.), the monitoring framework gathers information for Key Performance Indicators and software metrics. A collector component acquires the metrics and stores them in a database. Finally, monitoring information is accessible via TCloud monitoring API.

3.6.2.3 Quarter 7

Automatic Insertion of Monitoring Probes Monitoring software probes are inserted by the Service Manager inside the virtual machine by using the StratusLab contextualization mechanisms. This means that when the VM has been deployed, automatically it contains a set of probes that provide specific metrics for the applications to the monitoring system.

Integration of monitoring systems and Claudia for scalability Claudia is the service manager for managing the service lifecycle and the monitoring systems collects monitoring information of those services. The Key Performance Indicators (KPIs) that drive scalability are provided by probes and stored in the monitoring systems. Thus, a new process is required to obtain KPIs from the monitoring systems and inject them into Claudia. This process has been included as part as the

Claudia software.

3.6.2.4 Quarter 8

Placement decision module During this quarter a placement module has been created. This module is in charge of deciding the site where the service is going to be deployed. Thus, according a set of requirements specified in the service definition in the OVF, the placement module returns the site and its management URL. It is used for Cloud brokering.

Cloud Brokering with Claudia Cloud service brokering is a form of cloud service intermediation, in which an entity adds value to one or more cloud services on behalf of one or more consumers of that service. In StratusLab, Claudia is taking care of this role and is performing this functionality by using the placement decision module.

Integration of Claudia with OpenNebula for Hybrid Cloud Computing OpenNebula has been enhanced to provide the hybrid Cloud functionality. Claudia has been adapted to this functionality changing the way it obtains the IP where the virtual machine is deployed in a remote site.

Enhanced Claudia Works with Persistent Disk Service Claudia code has been enhanced to include the persistent disk functionality. References in the OVF it are translated to the appropriate fields in the OpenNebula machine template.

3.6.3 T6.2: Scalable and Elastic Management of Grid Site Infrastructure

3.6.3.1 Quarter 5

New Networking Model OpenNebula now provides an easily adaptable and customizable network subsystem in order to better integrate with the specific network requirements of existing data centers. The default configuration connects the virtual network interface to a bridge specified in the network template. Although this is flexible enough to fit into every datacenter it lacks two essential features: network isolation and firewall management. These features are provided as hooks which need to be activated and configured in OpenNebula's main configuration file.

3.6.3.2 Quarter 6

Support for User Data Injection in VMs A common requirement for VM contextualization is the ability to push user specific data into the VM, most notably access keys. This requires first the ability to store arbitrary data associated with each user, and then a flexible method to select and include user data in the context device. Both requirements are now met by OpenNebula.

Network Isolation through VLAN tagging OpenNebula provides support for host-managed VLANs to restrict network access through VLAN tagging. This mechanism is compliant with the IEEE 802.1Q standard, but it requires support

from the hardware switches. Also, OpenNebula allows administrators to restrict network access through VLAN tagging with Open vSwitch, a production quality, multilayer virtual switch.

Linux Bridge Firewall Management Alternatively to the use of VLAN tagging, it is possible to restrict network access through ebtables rules. The ebtables program enables transparent filtering of network traffic passing through a Linux bridge. This complements the previous automatic setup of simple firewall rules for TCP/UDP ports and ICMP traffic.

Accounting and monitoring The OpenNebula accounting module tracks information of the resource utilization that will be stored in predefined intervals of time. This information can be retrieved by User, Virtual Machine and Host. The OpenNebula statistics module keeps a predefined number of samples containing statistics for Hosts and Virtual Machines. These samples are built from the information that is retrieved for each resource by the OpenNebula Information Manager.

3.6.3.3 Quarter 7

Improved Network Management The network operations are now coupled with the VM lifecycle. This simplifies the management of networking (no hooks are needed) and solves previous issues with VLANs when migrating and restoring VMs. The network drivers define three actions (pre-boot, post-boot and clean) that can be easily customized if needed. Previous hooks for 802.1Q VLAN tagging, Open vSwitch and ebtables-based VLANs have been transformed to drivers, as well as the one to set simple firewalling rules.

Flexible Network Definition Networks can be now defined with an arbitrary range including an starting and ending IP address, network and network mask, or CIDR notation. It is possible also to define a network and a starting IP address to lease addresses.

New Network Lease Operations Network leases can now be put on hold to reserve them. This comes in handy when there are some IPs within the VLAN already assigned (e.g. .1 to the gateway). When a lease is put on hold, OpenNebula will not use it for a VM, until it is released.

Placement Policy Definition The placement of the VMs can now be defined on a VM-basis (restricted to oneadmin) or globally for the datacenter. This allows admins to set a global optimization policy to meet specific goals. The scheduling includes four predefined policies: packing, striping, load-aware, and custom.

3.6.3.4 Quarter 8

Support for Multiple Datastores The storage capabilities of OpenNebula have been improved in StratusLab with the addition of the Datastore abstraction. A datastore, previously known as an Image Repository, is an abstraction of any storage medium for VM disk images. Datastores are distributed to the hosts with specific transfer drivers. This allows a single host to include multiple datastores of dif-

ferent types. OpenNebula 3.4 includes four datastore types (system, file-system, iSCSI/LVM and VMware). The system has been architected to be highly modular, so these base types can be easily adapted to any specific deployment.

New Transfer Drivers Hosts are not tied to a single transfer mechanism (transfer driver) and now can access images from different datastores in different ways. Even, a VM can have its disks in different datastores. Also the transfers associated with persistent or `save_as` images have been simplified. There are also new drivers to use in combination with the datastores: qcow2, iSCSI and an improved version of vmware that uses the vmdkfs tools, which add to the shared and ssh drivers in OpenNebula 3.4.

Improved Amazon EC2 Adapter for Cloud Bursting OpenNebula 3.4 includes an improved driver to create hybrid clouds with Amazon EC2, to support most of the EC2 features like tags, security groups or VPC (Virtual Private Cluster).

3.6.4 T6.3: Cloud-like Interfaces Specific for the Scientific Community

3.6.4.1 Quarter 5

Drivers for accessing to Virtual Machine Managers Claudia accesses to the different Virtual Machine Managers (e.g. OpenNebula) by using the TCloud API. The implementation of this API has been done to work as an aggregated API, so that, it can be used to invoke different Cloud providers. In Y1, a driver for OpenNebula was created. In QR5, a driver to access to Flexiscale public Cloud has been incorporated in order to work towards an brokered federation.

Authentication in the Service Manager API In order to make possible the usage of Claudia in a public infrastructure, authentication mechanisms are being integrated in Claudia. A authentication proxy is being developed in WP2, and has been integrated in the claudia client used to access to Claudia.

Authorization Using Groups and Roles One of the main characteristics of an IaaS cloud is its multi-tenancy nature. In order to efficiently implement a multi-tenant system it is needed a flexible user system that allows the definition of user groups and access control lists to define specific access rights to each virtual resource. Therefore, the user system of OpenNebula has been extended to support groups. A user now is part of an user group, by default users in the same group can list and share (if labeled as public) any resource type (network, virtual machine, disk images etc.). The access control to each resource has been also improved with the addition of ACLs. An ACL express the user (or set of users) that may perform a given operation (e.g. create, delete or deploy) on a given virtual resource or set of them (e.g. virtual machines, networks, hosts or images).

VM Template Repository Usually IaaS clouds offer a predefined set of virtual machines (instance types) that users may instantiate. This leads to a simplification of the provisioning interface for final users that are only allowed to instantiate a

predefined error-free set of virtual machines. The predefined instances may include different OS types, packed with multiple software stacks (i.e. virtual appliances). This mechanism is even more robust in combination with the new group and user access control lists. In this way, access to a given instance type can be granted only to a specific set of users. Therefore, OpenNebula has been extended to include a new Template Repository that allows OpenNebula administrators and users to register virtual machine definitions in the system, to be instantiated later. These templates can be instantiated several times, and also shared with other users.

3.6.4.2 Quarter 6

Improvements in the Auth Module The OpenNebula authentication and authorization system has been extended in three areas. First, it now avoids some potential security holes when the end-user may choose the driver to authenticate with OpenNebula, specially when using X509 certificates. Second, the security of public cloud API has been strengthened by including special server accounts to run the services, these server processes can use symmetric cryptographic ciphers or X509 to authenticate with OpenNebula. And third, there have been improvements in the X509 and SSH authentication methods like native support for proxies or better support for DN strings.

3.6.4.3 Quarter 7

Introduction of TCloud API for monitoring API Monitoring information for both VMs and services can be accessible by a TCloud API. Thus, a TCloud driver for monitoring has been included as part of the implementation.

Security and User Management Some attributes (DISK/SOURCE, CONTEXT/FILES, NIC/MAC and NIC/VLAN_ID) have been restricted in VM Templates, because they can be easily used to gain oneadmin access or to comprise VMs of any user. There are new auth drivers for LDAP, with base and group filtering, and for Cloud API and OpenNebula front-end servers (server-based drivers). Also, as some of the drivers may take some time to authenticate a request (e.g. LDAP), session tokens can now be cached by OpenNebula. Finally, a new permission set has been included to manage access control to virtual resources. The new permissions overcomes the limitations of the previous PUBLIC attribute and allow users to share resources in multiple ways. Combined with the ACL system (also simplified to match the new permissions), this allows the implementation of multiple roles.

OCCI Extensions OpenNebula's OCCI API has been extended to include VM types, that can now be defined in the server configuration file and tagged with arbitrary information, like size, QoS parameters or price. These VM types can be programatically queried through the API. Also, to support the new VLAN features in OpenNebula, the OCCI networks can now be defined through a template, as for Virtual Machines.

Deltacloud API Updates Apache Deltacloud is a REST-based API that abstracts the differences between clouds, enabling the management of resources in different IaaS clouds using a single API. A series of back-end drivers translate the request to each cloud provider's native API. Currently, all major cloud service providers are supported.The OpenNebula back-end driver in Deltacloud has been updated to interact with OpenNebula 3.x clouds. This way, StratusLab sites could be accessed through the Deltacloud API, server and tools.

3.6.4.4 Quarter 8

OCCI updates The OCCI server has been improved to include user/group information in resources and extended information of resources.

New CloudAuth Driver This new driver delegates the authentication to the OpenNebula core. Therefore any OpenNebula auth driver can be used to authenticate cloud users or the Sunstone web UI.

Cloud Partitioning Cloud requests can be routed to an specific cluster with its own storage and network resources to better isolate public cloud users. Usually medium to large cloud sites structure their resources on multiple cluster each one with their own storage and networking systems. Cloud users are assigned to an specific cluster to prevent image trashing across large datacenters. This feature extends the previous cluster concept available in OpenNebula 3.0.

Improved Logging A new framework has been included to add logging information to the servers. In particular the logging facilities of the OpenNebula's Cloud API have been extended to ease the maintenance and deployment of several services like the Sunstone graphical interface or the OCCI API server (also used by the DeltaCloud API drivers).

ONE2ONE Driver for Cloud Federation OpenNebula manages external cloud as if they were local hosts. In this way the remote cloud is abstracted as a single but powerful resource. In order to interact with this special host a specific driver is needed that interacts with the remote cloud API in order to monitor and manage VM instances. In this period we have developed an driver that interacts with OpenNebula based clouds. It also integrates resources from other StratusLab services, specially the Marketplace.

3.6.5 Issues and Corrective Actions

3.6.5.1 Quarter 5

None.

3.6.5.2 Quarter 6

The document D6.4 was delayed and instead sent at the end of Q6. Some additional analysis and discussion have taken place to clarify and define the work to be done in WP6 in the coming months. Moreover, the document has been reviewed and rewritten in order to have a better quality.

3.6.5.3 Quarter 7

None.

3.6.5.4 Quarter 8

None.

4 Project Management

4.1 Consortium

The project consortium consisting of six partners (CNRS, UCM, GRNET, SIXSQ, TID, and TCD) has not changed since the start of the project. There have been no changes in the legal status of those partners; however, SixSq is now recognized as an SME by the European Commission. The representatives for TCD and TID have changed because of retirements and internal reorganization of activities.

4.2 Management Tasks

Meetings Tables 4.1–4.8 contain a list of the meetings by quarter that have been planned to foster collaboration between the project participants. Not listed are the planning meetings for each development sprint and the daily standup meetings.

Metrics Table 4.9 contains the metrics for the project. The table groups related metrics together.

Deliverables and Milestones Tables 5.1, 5.2, and 5.3 list all of the documents. In addition, these are available from the project website. The deliverables and other project documents (excluding milestones) have been uploaded to the HAL[1] open access repository to ensure their longterm availability.

4.3 Issues

No particular issues have arisen in Q8.

4.4 Planning

See the deliverable describing the sustainability and exploitation plan (D3.5) for information about the planning after the project.

[1] http://hal.archives-ouvertes.fr/

Table 4.1: *Meetings (Q1)*

Title	Date	Venue	Comments
StratusLab Kick-Off Meeting	14-15/06/2010	Orsay, FR	Kick-off of project. Detailed planning for accomplishing objectives. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1129
Technical Meeting	22/07/2010	Madrid, ES	Detailed technical discussions for StratusLab development. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1189
Sprint 1 Demo	30/07/2010	Phone/EVO	Sprint 1 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1191
Sprint 2 Demo	20/08/2010	Phone/EVO	Sprint 2 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1192

Table 4.2: *Meetings (Q2)*

Title	Date	Venue	Comments
Project Management Board	03/09/2010	Phone	PMB meeting to decide IPR policies. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1203
Sprint 3 Demo	10/09/2010	Phone/EVO	Sprint 3 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1203
Technical Meeting (TSCG)	21/09/2010	Phone/EVO	Shaping StratusLab distribution. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1213
WP6 research lines meeting	27/09/2010	Madrid, ES	Discussion about the main gaps identified in WP4 and some technologies to solve them. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1318
WP6 kickoff meeting	07/10/2010	Phone	Presentation of the lines to work on WP6 and distribution of work. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1320
Sprint 4 Demo	08/10/2010	Phone/EVO	Sprint 4 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1232
WP6 monitoring and accounting	26/10/2010	Phone	Audioconference about monitoring and accounting in StratusLab. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1321
Sprint 5 Demo	08/11/2010	Phone/EVO	Sprint 5 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1255
Face-to-Face Technical Meeting	15-16/11/2010	IBCP, Lyon, France	Discussion of StratusLab roadmap. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1243
Project Management Board	22/11/2010	Phone	Project overview; LoS policy. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1263

Table 4.3: *Meetings (Q3)*

Title	Date	Venue	Comments
Sprint 6 Demo	09/12/2010	Phone/EVO	Sprint 6 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1310
Sprint 7 Demo	17/12/2010	Phone/EVO	Sprint 7 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1323
Technical Meeting (TSCG)	27/01/2011	Phone/EVO	Feedback from EGI; priorities for distribution. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1213
Sprint 8 Demo	31/01/2011	Phone/EVO	Sprint 8 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1423
Technical Meeting (TSCG)	17/02/2011	Phone/EVO	Error reporting; priorities for next sprint. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1213
Sprint 9 Demo	18/02/2011	Phone/EVO	Sprint 9 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1442
Project Management Board	24/02/2011	Phone	Project status; MoUs; effort utilization; review planning. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1440

Table 4.4: *Meetings (Q4)*

Title	Date	Venue	Comments
Sprint 10 Demo	03/03/2011	Phone/EVO	Sprint 10 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1448
Technical Meeting (TSCG)	03/03/2011	Phone/EVO	Review of developments and priorities. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1460
Sprint 11 Demo	31/03/2011	Phone/EVO	Sprint 11 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1470
Metadata & Marketplace Demo	08/04/2011	EVO	Demo for HEPiX Virtualization Working Group. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1477
Sprint 12 Demo	29/04/2011	Phone/EVO	Sprint 12 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1492
Grid site deployment with Claudia (TID, GRNET)	09/05/2011	Phone	Discussion about how to use Claudia for the deployment of a grid site. http://indico2.lal.in2p3.fr/indico/conferenceTimeTable.py?confId=1530#20110509
Technical Meeting (TSCG)	10/05/2011	Phone	Persistent storage and cloud interfaces. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1526
Interproject Collaboration	11/05/2011	Amsterdam	StratusLab, HPC Cloud, and Mantychore discussions. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1510
Sprint 13 Demo	16/05/2011	Phone/EVO	Sprint 13 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1513
Integration Meeting	23-24/05/2011	Geneva	F2F meeting for 1.0 release. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1503
Interproject Collaboration	27/05/2011	Phone	Discussion with Contrail project. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1527
Grid site deployment and scalability (TID, GRNET)	27/05/2011	Phone	Discussion to align the work. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1529

Table 4.5: *Meetings (Q5)*

Title	Date	Venue	Comments
Sprint 14 Demo	10/06/2011	Phone/EVO	Sprint 14 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1533
StratusLab Face-to-Face Meeting	21-23/06/2011	Geneva, CH	Integration of software. Update of roadmap. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1502
Sprint 15 Demo	23/06/2011	Phone/EVO	Sprint 15 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1541
StratusLab First Periodic Review	04/07/2011	Brussels, BE	External review of project's progress. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1501
Sprint 16 Demo	29/07/2011	Phone/EVO	Sprint 16 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1557
Technical Meeting (TSCG)	25/08/2011	Phone/EVO	Priorities for upcoming sprints. Architecture for StratusLab 2.0. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1560
Sprint 2 Demo	20/08/2010	Phone/EVO	Sprint 2 demonstration meeting. http://indico.lal.in2p3.fr/conferenceDisplay.py?confId=1192

Table 4.6: *Meetings (Q6)*

Title	Date	Venue	Comments
Project Management Board	12/09/2011	Phone	Project status; Reviewer Feedback; Y2 Effort and Budgets. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1561
Sprint 17 Demo	16/09/2011	Phone/EVO	Sprint 17 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1574
Technical Meeting (TSCG)	26/09/2011	Phone/EVO	Architecture review; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1575
Technical Meeting (TSCG)	17/10/2011	Phone/EVO	Use Cases for Y2 Review; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1613
Sprint 18 Demo	19/10/2011	Phone/EVO	Sprint 18 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1617
Sprint 19 Demo	16/11/2011	Phone/EVO	Sprint 19 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1638
StratusLab Cloud Training	17–18/11/2011	Orsay, France	General training for users and administrators. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1565
Technical Meeting (TSCG)	21/11/2011	Phone/EVO	F2F agenda; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1642
Project Management Board	30/11/2011	Phone	Project status; ERINA+ MoU; Reviewer Recommendations; Sustainability. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1650
StratusLab Face-to-Face Meeting	30/11/2011–02/12/2011	Dublin, Ireland	Integration of software; Update of roadmap. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1621

Table 4.7: *Meetings (Q7)*

Title	Date	Venue	Comments
Sprint 20 Demo	19/12/2011	Phone/EVO	Sprint 20 demonstration meeting. http://indico2.lal.in2p3.fr/indico/categoryDisplay.py?categId=131
Technical Meeting (TSCG)	31/01/2012	Phone/EVO	F2F agenda; priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1709
Sprint 21 Demo	30/01/2012	Phone/EVO	Sprint 21 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1705
Sprint 22 Demo	28/02/2012	Phone/EVO	Sprint 22 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1749
Technical Meeting (TSCG)	01/03/2012	Phone/EVO	Priorities for upcoming sprints. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1709

Table 4.8: *Meetings (Q8)*

Title	Date	Venue	Comments
Sprint 23 Demo	22/03/2012	Phone/EVO	Sprint 23 demonstration meeting. http://indico2.lal.in2p3.fr/indico/categoryDisplay.py?categId=1779
Technical Meeting (TSCG)	18/04/2012	Athens, Greece	Planning for EU Review. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1770
Project Management Board (PMB)	18/04/2012	Athens, Greece	Planning for EU Review; financial review. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1770
Face to Face Meeting	18-20/04/2012	Athens, Greece	StratusLab Face to Face Meeting http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1770
Sprint 24 Demo	27/04/2012	Phone/EVO	Sprint 24 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1828
Sprint 25 Demo	??/05/2012	Phone/EVO	21 stories demonstrated; see JIRA for details.
Sprint 26 Demo	29/05/2012	Phone/EVO	Sprint 26 demonstration meeting. http://indico2.lal.in2p3.fr/indico/conferenceDisplay.py?confId=1845

Table 4.9: *Metrics*

Metric	Q2	Q3	Q4	Y1 Target	Q5	Q6	Q7	Q8	Y2 Target
No. of people trained on StratusLab software	N/A	N/A	~ 25	–	~36	~89	0	25	–
No. of people on StratusLab announcement list	67	67	67	25	70	72	72	73	75
Registered users on StratusLab discussion site	N/A	N/A	N/A	50	N/A	12	13	16	100
No. of views of website	2922	4623	4579	–	5472	4985	4442	5372	–
No. of completed sprints	5	5	4	–	3	3	3	4	–
No. of releases	1	1	1	–	2	1	2	2	–
No. of open user stories	38	72	101	–	118	107	121	102	–
No. of implemented user stories	69	40	67	–	50	48	57	61	–
No. of open bugs	6	15	22	–	28	51	51	66	–
No. of fixed bugs	7	11	27	–	14	20	17	14	–
No. of prod. sites running StratusLab dist.	1	1	1	5	1	3	5		10
Availability of hosted grid sites	N/A	N/A	100%	80%	91%	74%	93%	98%	95%
Reliability of hosted grid sites	N/A	N/A	100%	80%	92%	78%	93%	98%	95%
No. of VOs served via StratusLab hosted grid sites	0	1	1	10	21	18	18	18	30
No. of sci. disciplines served via StratusLab hosted grid sites	0	0	0	3	11	9	9	9	15
Delivered computing resources through hosted grid services	N/A	16 cores	16 cores	–	32 cores	32 cores	32 cores	32 cores	–
Delivered computing resources through hosted cloud services	N/A	256 cores	256 cores	–	256 cores	288 cores	256 cores	256 core	–
Storage provided through cloud service	N/A	N/A	N/A	–	0	0	0	3 TB	3 TB
No. of jobs run in hosted grid site	N/A	N/A	N/A	–	13,960	16,916	28,345	12,930	–
Norm. CPU time consumed in the hosted grid site (hrs)	N/A	N/A	N/A	–	26,202	14,231	87,671	51,895	–
No. base machine images	5	7	8	5	8	13	–	–	10
No. of base machine image downloads	783	2628	7072	–	7225	6657	–	–	–
No. appliances	0	6	7	5	7	7	–	–	15
No. of appliance downloads	0	252	687	–	1010	426	–	–	–
No. of Marketplace metadata entries	–	–	–	–	–	–	111	114	–
No. of Marketplace endorsers	–	–	–	–	–	–	24	35	–
No. of Marketplace base images	–	–	–	–	–	–	86	71	–
No. of Marketplace appliances	–	–	–	–	–	–	25	43	–

5 Deliverables and Milestones

Tables 5.1 and 5.2 show the deliverables for the first and second years of the project. Table 5.3 lists all of the milestones. All of the deliverables and milestones for the first year of the project have been produced and submitted as foreseen in the project's roadmap. All of these are available from the project's website[1].

Two technical notes have also been produced during the first year: "StratusLab Marketplace" describing the technical specification of the Marketplace and "Installing and operating a production grid site in the StratusLab cloud: Experience and issues" providing feedback to developers and advice to administrators running grid services within a cloud. These notes are also available from the project website.

[1] http://stratuslab.eu/doku.php/deliverables

Table 5.1: *Deliverables (Year 1)*

No.	Title	Version	WP No.	Lead Beneficiary	Nature	Diss. Level	Due Date	Actual Date	Status	Contractual	Comments
D2.1	Review of the Use of Cloud and Virtualization Technologies in Grid Infrastructures	1.2	WP2	CNRS	R	PU	PM2	11/08/2010	Done	Yes	
D4.1	Reference Architecture for StratusLab Toolkit 1.0	1.0	WP4	SIXSQ	R	PU	PM3	14/09/2010	Done	Yes	
D5.1	Infrastructure Specification	1.0	WP5	GRNET	R	PU	PM3	14/09/2010	Done	Yes	
D3.1	Initial Plan for Dissemination, Collaboration and Standardization Activities	1.0	WP3	TCD	R	PU	PM4	18/10/2010	Done	Yes	
D6.1	Cloud-like Management of Grid Sites 1.0 Design Report	1.0	WP6	TID	R	PU	PM5	16/11/2010	Done	Yes	
D5.2	Infrastructure Tool and Policy Specification	1.0	WP5	GRNET	R	PU	PM6	15/12/2010	Done	Yes	
D6.2	Cloud-like Management of Grid Sites 1.0 Software	1.1	WP6	TID	P	PU	PM11	13/05/2011	Done	Yes	
D2.2	Report on Evaluation of StratusLab Products	1.0	WP2	CNRS	R	PU	PM12	15/06/2011	Done	Yes	
D3.2	Report on Dissemination, Collaboration and Standardization Activities	1.1	WP3	TCD	R	PU	PM12	16/06/2011	Done	Yes	
D3.3	Exploitation and Sustainability First Plan	1.1	WP3	TCD	R	PU	PM12	16/06/2011	Done	Yes	
D4.2	StratusLab Toolkit 1.0	1.0	WP4	SIXSQ	P	PU	PM12	15/06/2011	Done	Yes	
D4.3	First Year Software Integration Report	1.0	WP4	SIXSQ	R	PU	PM12	15/06/2011	Done	Yes	
D5.3	First Year Infrastructure Operations Report	1.1	WP5	GRNET	R	PU	PM12	16/06/2011	Done	Yes	
D6.3	First Year Cloud-like Management of Grid Sites Research Report	1.0	WP6	TID	R	PU	PM12	15/06/2011	Done	Yes	

Table 5.2: *Deliverables (Year 2)*

No.	Title	Version	WP No.	Lead Beneficiary	Nature	Diss. Level	Due Date	Actual Date	Status	Contractual	Comments
D2.3	Survey of Targeted Communities Concerning StratusLab		WP2	CNRS	R	PU	PM14	12/08/2011	Done	Yes	
D4.4	Reference Architecture for StratusLab Toolkit 2.0		WP4	SIXSQ	R	PU	PM15	03/10/2011	Done	Yes	Delayed PM15-16
D6.4	Cloud-like Management of Grid Sites 2.0 Design Report		WP6	TID	R	PU	PM17	15/12/2011	Done	Yes	Delayed PM17-18
D5.4	Economic Analysis of Infrastructure Operations		WP5	GRNET	R	PU	PM18	19/03/2012	Done	Yes	Delayed to PM21
D6.5	Cloud-like Management of Grid Sites 2.0 Software		WP6	TID	P	PU	PM23	11/05/2012	Done	Yes	
D2.4	Final Report on StratusLab Adoption		WP2	CNRS	R	PU	PM24	06/06/2012	Done	Yes	
D2.5	Report on Evaluation of StratusLab Products		WP2	CNRS	R	PU	PM24	14/06/2012	Done	Yes	
D3.4	Final Review of Dissemination, Collaboration and Standardization Activities		WP3	TCD	R	PU	PM24	29/05/2012	Done	Yes	
D3.5	Exploitation and Sustainability Final Plan		WP3	TCD	R	PU	PM24	22/05/2012	Done	Yes	
D4.5	StratusLab Toolkit 2.0		WP4	SIXSQ	P	PU	PM24	28/05/2012	Done	Yes	
D4.6	Software Integration Final Report		WP4	SIXSQ	R	PU	PM24	28/05/2012	Done	Yes	
D5.5	Infrastructure Operations Final Report		WP5	GRNET	R	PU	PM24	05/06/2012	Done	Yes	
D6.6	Cloud-like Management of Grid Sites Research Final Report		WP6	TID	R	PU	PM24	04/06/2012	Done	Yes	

Table 5.3: *Milestones*

No.	Title	WP No.	Lead Beneficiary	Due Date	Achieved	Actual Date	Comments
MS1	Establishment of Management Infrastructure and Metrics Definition	WP1	CNRS	PM3	Yes	1/09/2010	
MS6	Website Operational	WP3	TCD	PM3	Yes	6/09/2010	
MS2	Contact Procedures and Supporting Tools for Targeted Communities	WP2	CNRS	PM4	Yes	10/12/2010	
MS7	StratusLab Development, Certification and Release Procedures in Place	WP4	SIXSQ	PM6	Yes	10/12/2010	
MS3	Creation of Virtual Appliances for Bioinformatics Community	WP2	CNRS	PM9	Yes	14/03/2011	
MS10	Initial virtual appliance repository	WP5	GRNET	PM9	Yes	4/03/2011	
MS14	Release of Cloud-like Management of Grid Services and Resources 1.0 Beta	WP6	TID	PM9	Yes	14/03/2011	
MS8	Release of StratusLab 1.0 Beta	WP4	SIXSQ	PM10	Yes	05/04/2011	
MS11	Operation of Site Running StratusLab toolkit v1.0	WP5	GRNET	PM10	Yes	04/04/2011	
MS4	Adoption of StratusLab Software by External Grid Sites	WP2	CNRS	PM14	Yes	19/03/2012	
MS12	Delivery of Virtual Appliance Repository	WP5	GRNET	PM18	Yes	15/12/2011	
MS5	Opening of Virtual Appliances Repository to External Application Communities	WP2	CNRS	PM20	Yes	19/03/2012	
MS15	Release of Cloud-like Management of Grid Services and Resources 2.0 Beta	WP6	TID	PM21	Yes	19/03/2012	
MS9	Release of StratusLab 2.0 Beta	WP4	SIXSQ	PM22	Yes	02/05/2012	
MS13	Operation of Site Running StratusLab Toolkit v2.0	WP5	GRNET	PM22	Yes	02/05/2012	

6 Use of Resources

See the “Use of Resources” annex for all of the financial information.

Glossary

APEL	Accounting Processor for Event Logs (EGI accounting tool)
Appliance	Virtual machine containing preconfigured software or services
CDMI	Cloud Data Management Interface (from SNIA)
CE	Computing Element in EGI
DCI	Distributed Computing Infrastructure
DMTF	Distributed Management Task Force
EGEE	Enabling Grids for E-sciencE
EGI	European Grid Infrastructure
EGI-TF	EGI Technical Forum
GPFS	General Parallel File System by IBM
Hybrid Cloud	Cloud infrastructure that federates resources between organizations
IaaS	Infrastructure as a Service
iSGTW	International Science Grid This Week
KPI	Key Performance Indicator
LB	Load Balancer
LRMS	Local Resource Management System
MoU	Memorandum of Understanding
NFS	Network File System
NGI	National Grid Initiative
OCCI	Open Cloud Computing Interface
OVF	Open Virtualization Format
Public Cloud	Cloud infrastructure accessible to people outside of the provider's organization
Private Cloud	Cloud infrastructure accessible only to the provider's users
SE	Storage Element in EGI
SGE	Sun Grid Engine
SNIA	Storage Networking Industry Association
TCloud	Cloud API based on vCloud API from VMware
VM	Virtual Machine
VO	Virtual Organization
VOBOX	Grid element that permits VO-specific service to run at a resource center
Worker Node	Grid node on which jobs are executed

XMLRPC	XML-based Remote Procedure Call
YAIM	YAIM Ain't an Installation Manager (configuration utility for EGI)